Random sampling locations for comparing a mean with a fixed threshold (parametric)

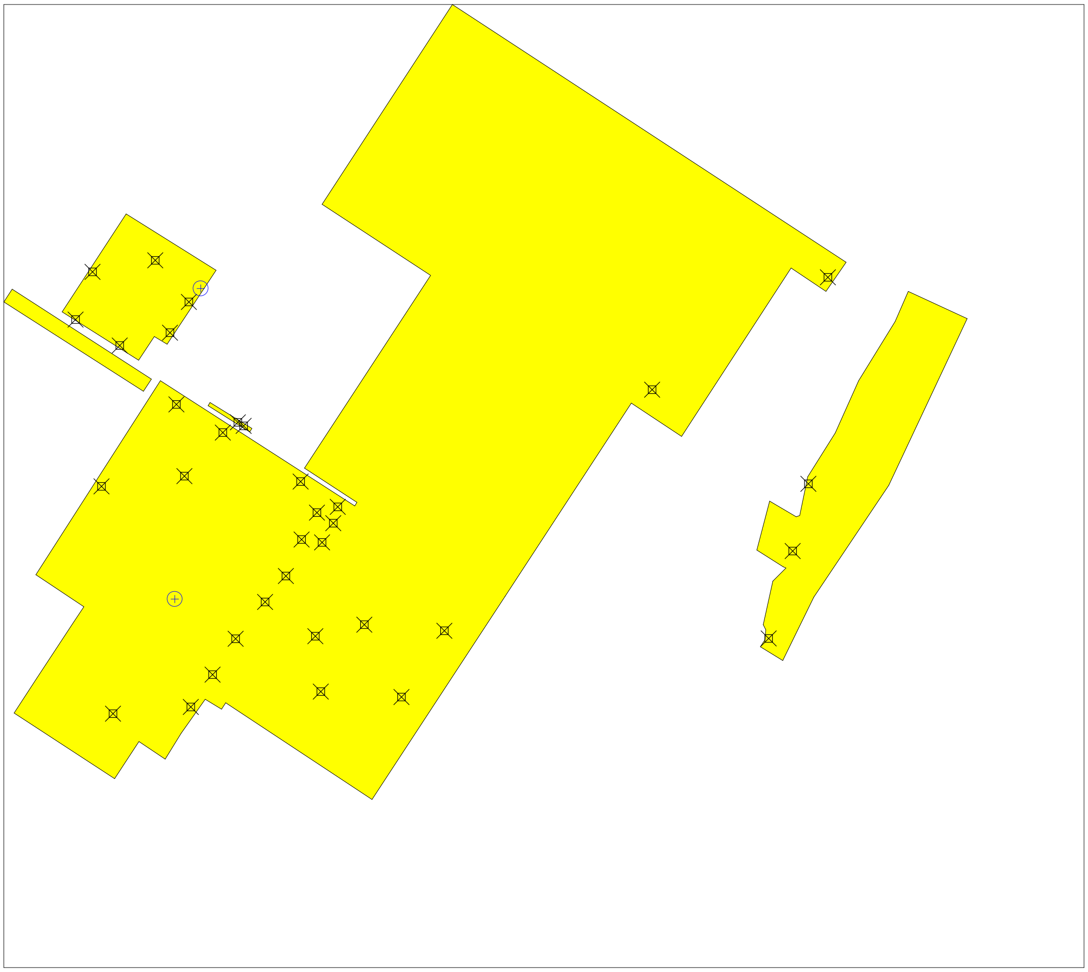
Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	36
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.

9558508



Area: Area 1					
X Coord	Y Coord	Label	Value	Type	Historical
679530.9430	3082575.4480	G-36SD	0.007	Manual	T
679606.6840	3082692.3830	G-37SD	0.0029	Manual	T
679671.3170	3082565.9250	G-46SD	0.0027	Manual	T
679745.9820	3082681.3860	G-47SD	0.0027	Manual	T
679532.9930	3082835.5820	J-42SD	0.0085	Manual	T
679552.9590	3082868.6600	J-43SD	0.0076	Manual	T
679521.7790	3082672.0220	J-54SD	0.0027	Manual	T
680108.0130	3083101.3520	J-57SD	0.0106	Manual	T
680413.8310	3083297.0130	J-58SD	0.0027	Manual	T
679149.4920	3082933.0980	TW01-13	0.0027	Manual	T
679279.7760	3083075.6320	TW01-14	0.0027	Manual	T
679293.5600	3082950.4980	TW01-17	0.0046	Manual	T
679360.5700	3083026.4980	TW01-18	0.0211	Manual	T
679169.0760	3082537.3510	TW01-27	0.01665	Manual	T
679495.8840	3082940.9730	TW01-33	0.0095	Manual	T
679304.6530	3082548.6880	TW01-34	0.0082	Manual	T

679342.7410	3082605.3190	TW01-35	0.0195	Manual	T
679382.8900	3082667.5270	TW01-36	0.0437	Manual	T
679433.9450	3082731.6820	TW01-37	0.0066	Manual	T
679470.3570	3082776.7350	TW01-38	0.0105	Manual	T
679497.3310	3082840.3960	TW01-39	0.0027	Manual	T
679524.3310	3082886.8990	TW01-40	0.0027	Manual	T
679560.6110	3082897.2580	TW01-41	0.0075	Manual	T
679276.8911	3082736.8178	J-59SD	0.0027	Random	

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680310.3290	3082668.1710	J-59SD	0.0027	Manual	T
680352.2560	3082820.3630	J-60SD	0.0027	Manual	T
680379.4090	3082937.1350	J-61SD	0.0027	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679386.3850	3083044.5490	TW06-63	0.0031	Manual	T
679396.8510	3083038.0640	TW06-64	0.0027	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**

Area: Area 5					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679133.4290	3083306.3130	TW01-01	0.0027	Manual	T
679104.2450	3083223.2620	TW01-02	0.0028	Manual	T
679242.7260	3083326.5280	TW01-07	0.0027	Manual	T
679181.2750	3083178.2880	TW01-08	0.0038	Manual	T
679268.7700	3083200.3260	TW01-11	0.0035	Manual	T
679301.1600	3083254.0340	TW01-12	0.0027	Manual	T
679322.2873	3083277.3101		0	Random	

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric

approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

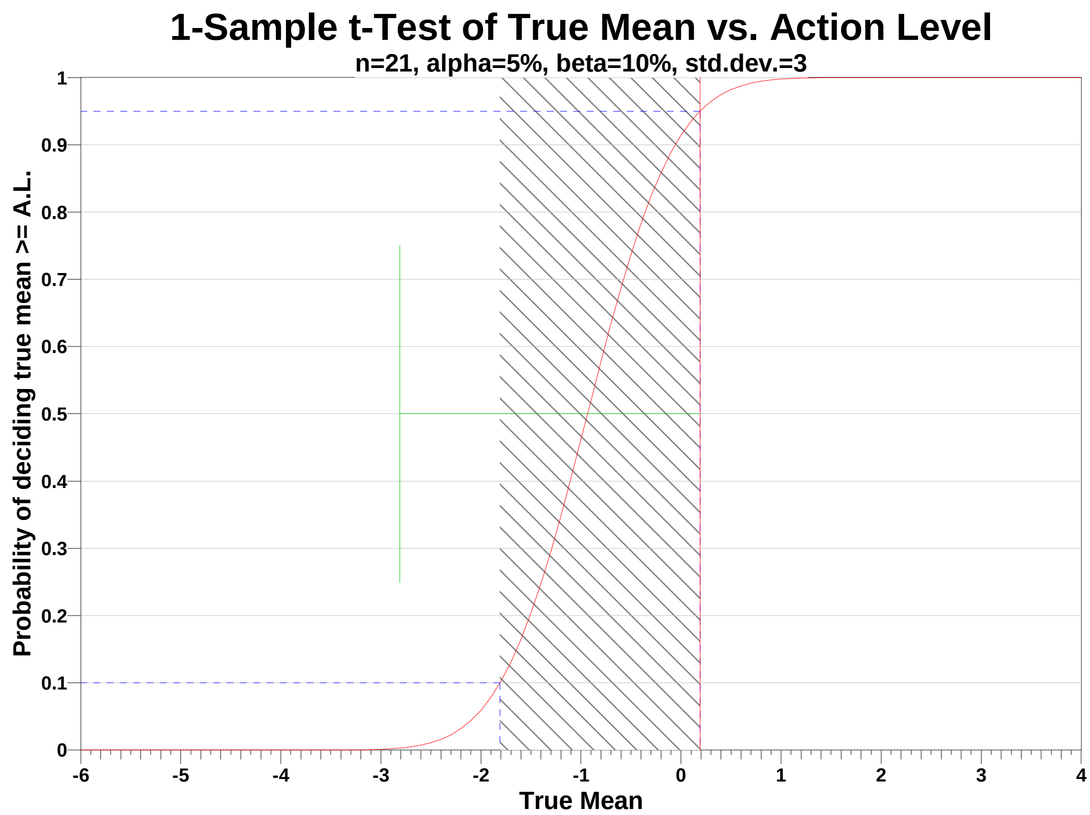
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.19		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	1079221	269807	854014	213504	716940	179236
	β=10	854014	213505	655131	163784	535818	133955
	β=15	716940	179236	535818	133956	428488	107123
LBGR=80	β=5	269807	67453	213504	53377	179236	44810
	β=10	213505	53378	163784	40947	133955	33490
	β=15	179236	44811	133956	33490	107123	26782
LBGR=70	β=5	119915	29980	94892	23724	79661	19916

	β=10	94892	23724	72794	18199	59536	14885
	β=15	79662	19917	59537	14885	47611	11903

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis
The following data points were entered by the user for analysis.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
0	0	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027
10	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0028	0.0029
20	0.0031	0.0031	0.0035	0.0038	0.0046	0.0066	0.007	0.0075	0.0076	0.0082
30	0.0085	0.0095	0.0105	0.0106	0.01665	0.0195	0.0211	0.0437		

SUMMARY STATISTICS	
n	38
Min	0
Max	0.0437
Range	0.0437
Mean	0.0064908
Median	0.00285
Variance	6.1556e-005
StdDev	0.0078458
Std Error	0.0012728
Skewness	3.3454
Interquartile Range	0.00505
Percentiles	

1%	5%	10%	25%	50%	75%	90%	95%	99%
0	0.002565	0.0027	0.0027	0.00285	0.00775	0.01694	0.02223	0.0437

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	4.743	3.01	Yes

The test statistic 4.743 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	0.0437

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Shapiro-Wilk test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Shapiro-Wilk Test Statistic	0.7026
Shapiro-Wilk 5% Critical Value	0.936

The calculated Shapiro-Wilk test statistic is less than the 5% Shapiro-Wilk critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

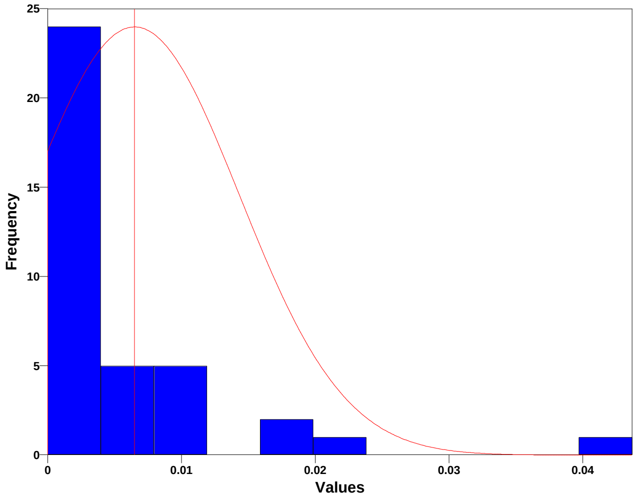
Data Plots
Graphical displays of the data are shown below.

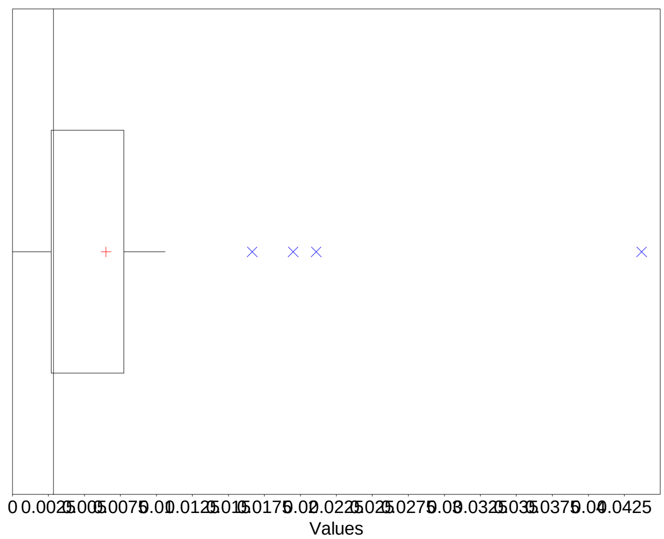
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

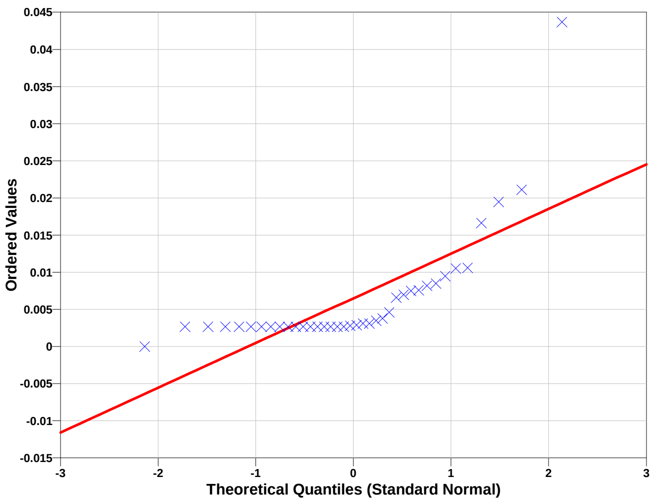
The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25th and 75th percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted

individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_n, for which a fraction p of the distribution is less than x_n. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Shapiro-Wilk (SW) test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Shapiro-Wilk Test Statistic	0.5919
Shapiro-Wilk 5% Critical Value	0.938

The calculated SW test statistic is less than the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance.

The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.008638
95% Non-Parametric (Chebyshev) UCL	0.01204

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.01204) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value t was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=38 data,
AL is the action level or threshold (0.19),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=37 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-144.18	1.6871	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
38	24	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

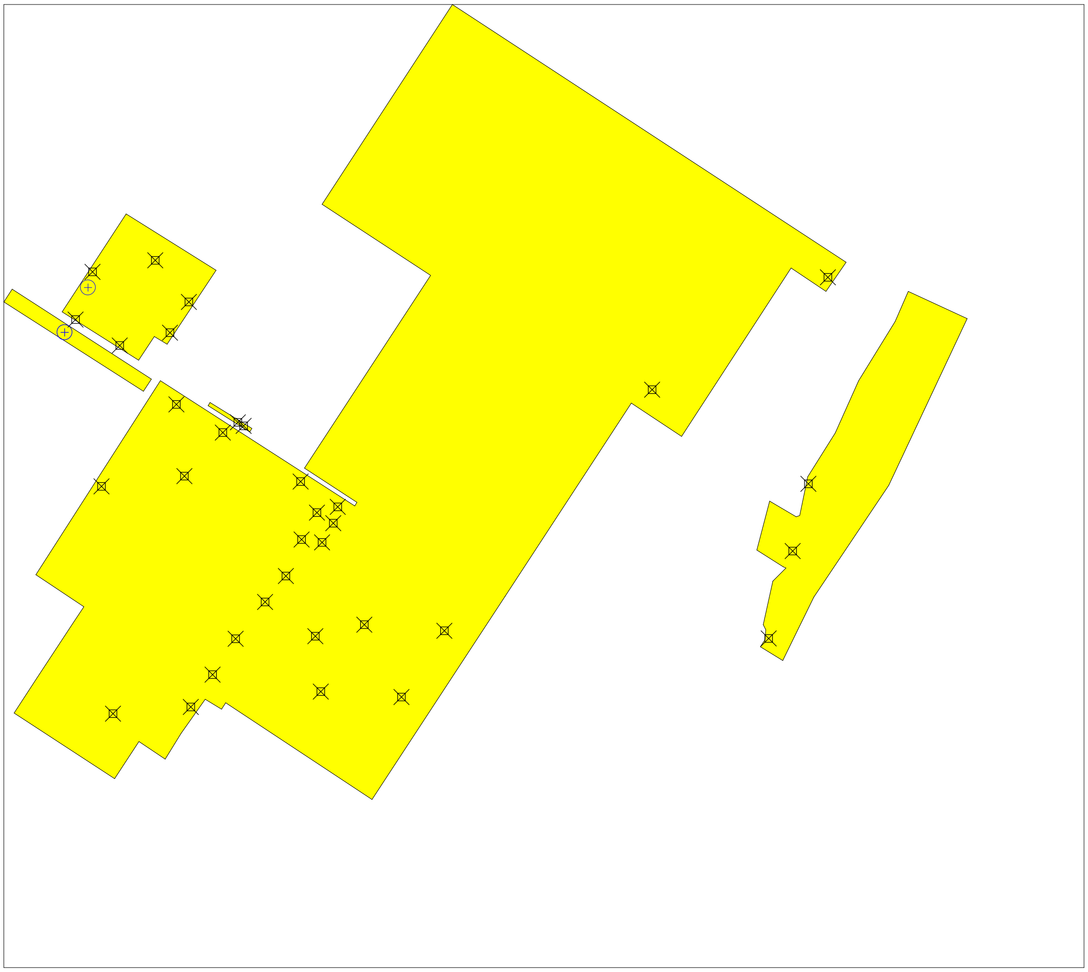
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	36
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679530.9430	3082575.4480	G-36SD	0.007	Manual	T
679606.6840	3082692.3830	G-37SD	0.0029	Manual	T
679671.3170	3082565.9250	G-46SD	0.0027	Manual	T
679745.9820	3082681.3860	G-47SD	0.0027	Manual	T
679532.9930	3082835.5820	J-42SD	0.0085	Manual	T
679552.9590	3082868.6600	J-43SD	0.0076	Manual	T
679521.7790	3082672.0220	J-54SD	0.0027	Manual	T
680108.0130	3083101.3520	J-57SD	0.0106	Manual	T
680413.8310	3083297.0130	J-58SD	0.0027	Manual	T
679149.4920	3082933.0980	TW01-13	0.0027	Manual	T
679279.7760	3083075.6320	TW01-14	0.0027	Manual	T
679293.5600	3082950.4980	TW01-17	0.0046	Manual	T
679360.5700	3083026.4980	TW01-18	0.0211	Manual	T
679169.0760	3082537.3510	TW01-27	0.01665	Manual	T
679495.8840	3082940.9730	TW01-33	0.0095	Manual	T
679304.6530	3082548.6880	TW01-34	0.0082	Manual	T

679342.7410	3082605.3190	TW01-35	0.0195	Manual	T
679382.8900	3082667.5270	TW01-36	0.0437	Manual	T
679433.9450	3082731.6820	TW01-37	0.0066	Manual	T
679470.3570	3082776.7350	TW01-38	0.0105	Manual	T
679497.3310	3082840.3960	TW01-39	0.0027	Manual	T
679524.3310	3082886.8990	TW01-40	0.0027	Manual	T
679560.6110	3082897.2580	TW01-41	0.0075	Manual	T

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680310.3290	3082668.1710	J-59SD	0.0027	Manual	T
680352.2560	3082820.3630	J-60SD	0.0027	Manual	T
680379.4090	3082937.1350	J-61SD	0.0027	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679386.3850	3083044.5490	TW06-63	0.0031	Manual	T
679396.8510	3083038.0640	TW06-64	0.0027	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679084.9690	3083201.1536	TW01-01	0.0027	Random	

Area: Area 5					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679133.4290	3083306.3130	TW01-01	0.0027	Manual	T
679104.2450	3083223.2620	TW01-02	0.0028	Manual	T
679242.7260	3083326.5280	TW01-07	0.0027	Manual	T
679181.2750	3083178.2880	TW01-08	0.0038	Manual	T
679268.7700	3083200.3260	TW01-11	0.0035	Manual	T
679301.1600	3083254.0340	TW01-12	0.0027	Manual	T
679126.0114	3083278.9929		0	Random	

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric

approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs

The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n	is the number of samples,
S	is the estimated standard deviation of the measured values including analytical error,
Δ	is the width of the gray region,
α	is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β	is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$	is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$	is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

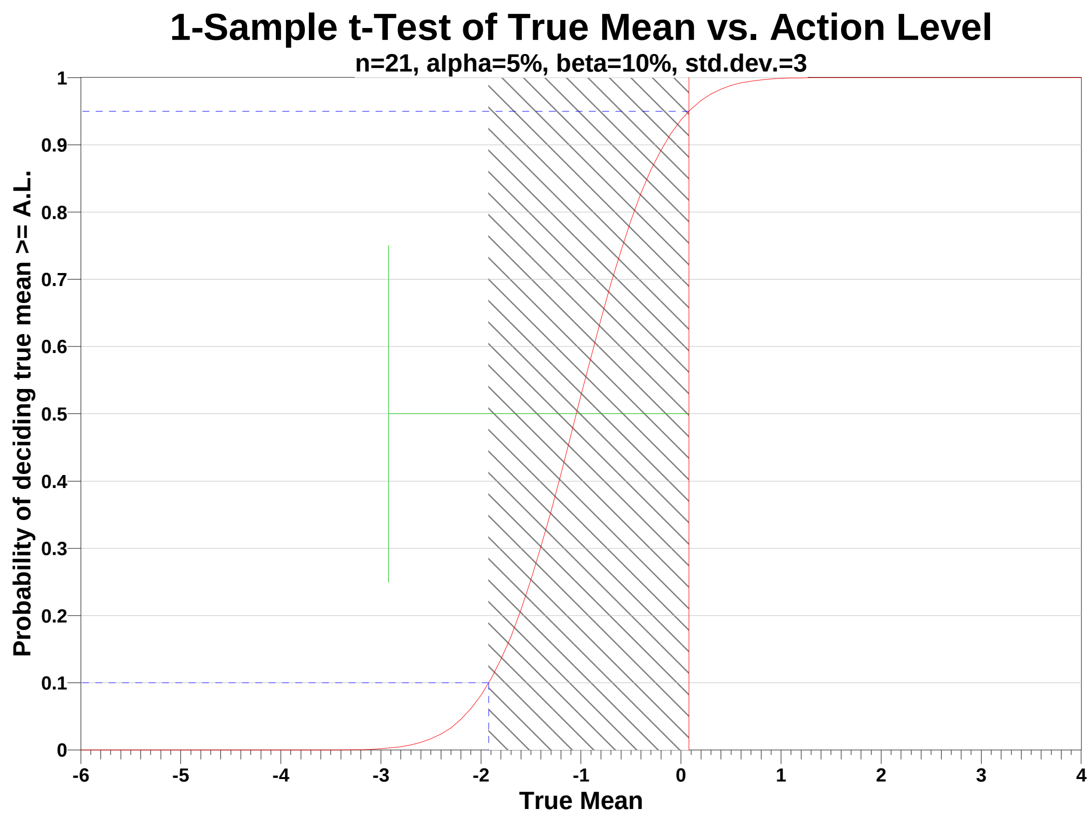
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.078		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	6403655	1600915	5067367	1266843	4254025	1063507
	β=10	5067367	1266843	3887278	971821	3179323	794832
	β=15	4254026	1063508	3179323	794832	2542472	635619
LBGR=80	β=5	1600915	400230	1266843	316712	1063507	265878
	β=10	1266843	316712	971821	242956	794832	198709
	β=15	1063508	265878	794832	198709	635619	158906
LBGR=70	β=5	711519	177881	563042	140761	472670	118168

	β=10	563042	140762	431921	107981	353259	88316
	β=15	472671	118169	353259	88316	282498	70625

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis
The following data points were entered by the user for analysis.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
0	0	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027
10	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0028	0.0029
20	0.0031	0.0031	0.0035	0.0038	0.0046	0.0066	0.007	0.0075	0.0076	0.0082
30	0.0085	0.0095	0.0105	0.0106	0.01665	0.0195	0.0211	0.0437		

SUMMARY STATISTICS	
n	38
Min	0
Max	0.0437
Range	0.0437
Mean	0.0064908
Median	0.00285
Variance	6.1556e-005
StdDev	0.0078458
Std Error	0.0012728
Skewness	3.3454
Interquartile Range	0.00505
Percentiles	

1%	5%	10%	25%	50%	75%	90%	95%	99%
0	0.002565	0.0027	0.0027	0.00285	0.00775	0.01694	0.02223	0.0437

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	4.743	3.01	Yes

The test statistic 4.743 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	0.0437

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Shapiro-Wilk test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Shapiro-Wilk Test Statistic	0.7026
Shapiro-Wilk 5% Critical Value	0.936

The calculated Shapiro-Wilk test statistic is less than the 5% Shapiro-Wilk critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

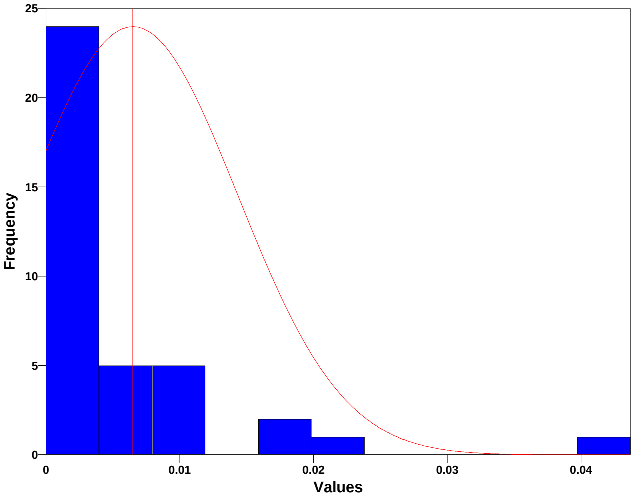
Data Plots
Graphical displays of the data are shown below.

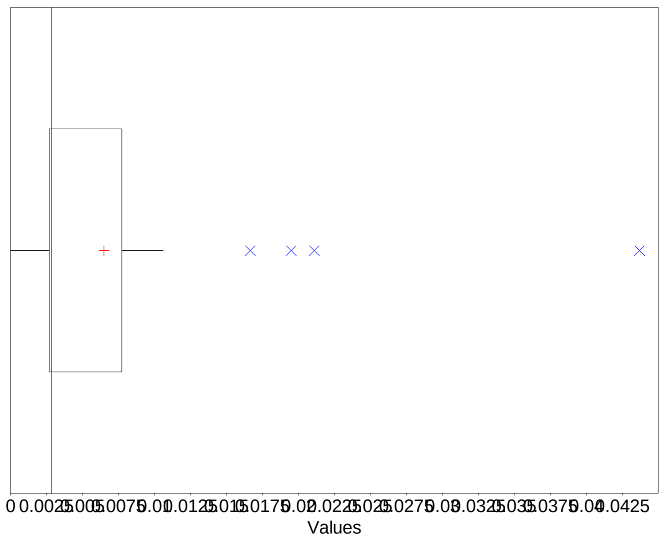
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

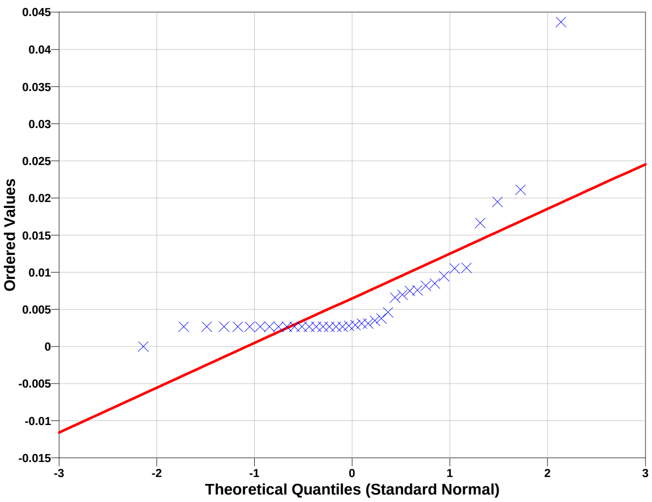
The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted

individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Shapiro-Wilk (SW) test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Shapiro-Wilk Test Statistic	0.5919
Shapiro-Wilk 5% Critical Value	0.938

The calculated SW test statistic is less than the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance.

The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.008638
95% Non-Parametric (Chebyshev) UCL	0.01204

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.01204) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test

A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where

x is the sample mean of the n=38 data,
AL is the action level or threshold (0.078),
SE is the standard error = (standard deviation) / (square root of n).

This *t* was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=37 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-56.185	1.6871	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
38	24	Reject

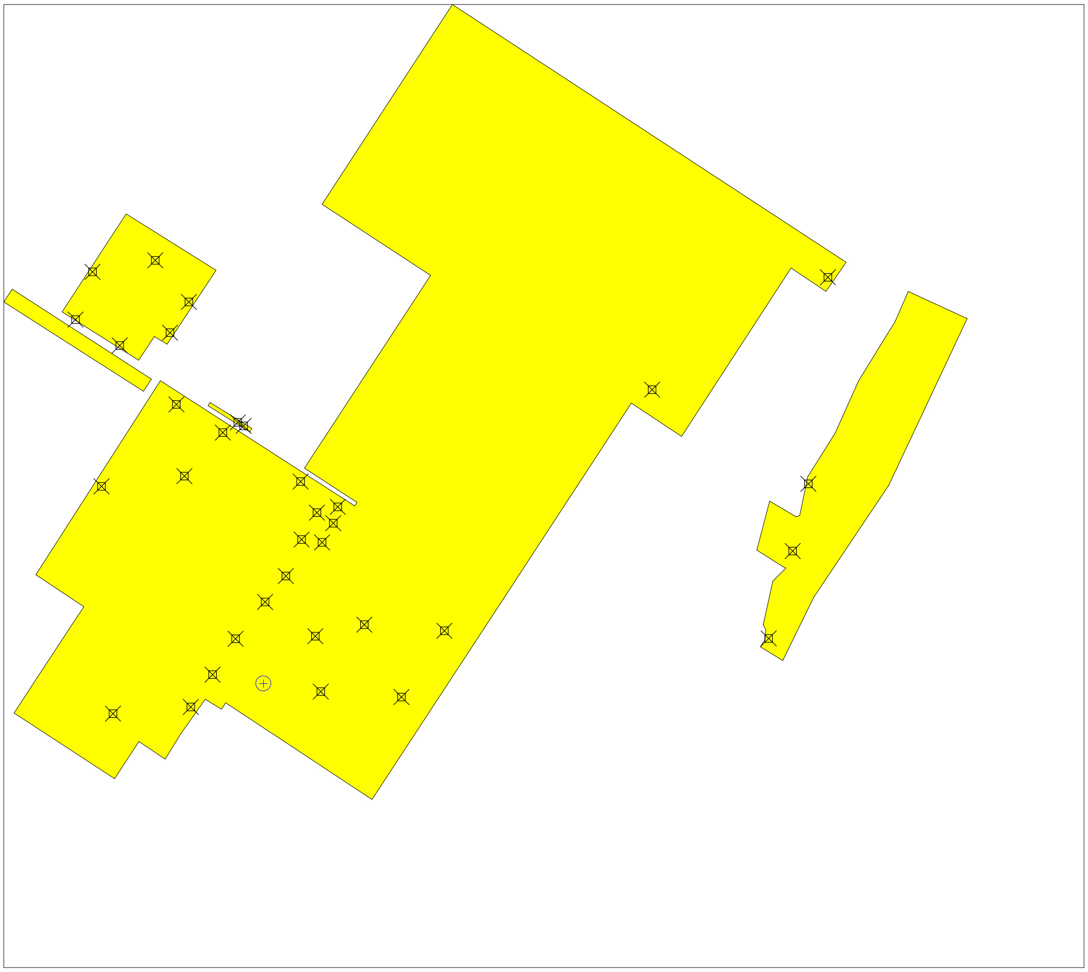
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	2
Number of samples on map [a]	35
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$2,000.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	Y Coord	Label	Value	Type	Historical
679530.9430	3082575.4480	G-36SD	0.007	Manual	T
679606.6840	3082692.3830	G-37SD	0.0029	Manual	T
679671.3170	3082565.9250	G-46SD	0.0027	Manual	T
679745.9820	3082681.3860	G-47SD	0.0027	Manual	T
679532.9930	3082835.5820	J-42SD	0.0085	Manual	T
679552.9590	3082868.6600	J-43SD	0.0076	Manual	T
679521.7790	3082672.0220	J-54SD	0.0027	Manual	T
680108.0130	3083101.3520	J-57SD	0.0106	Manual	T
680413.8310	3083297.0130	J-58SD	0.0027	Manual	T
679149.4920	3082933.0980	TW01-13	0.0027	Manual	T
679279.7760	3083075.6320	TW01-14	0.0027	Manual	T
679293.5600	3082950.4980	TW01-17	0.0046	Manual	T
679360.5700	3083026.4980	TW01-18	0.0211	Manual	T
679169.0760	3082537.3510	TW01-27	0.01665	Manual	T
679495.8840	3082940.9730	TW01-33	0.0095	Manual	T
679304.6530	3082548.6880	TW01-34	0.0082	Manual	T

679342.7410	3082605.3190	TW01-35	0.0195	Manual	T
679382.8900	3082667.5270	TW01-36	0.0437	Manual	T
679433.9450	3082731.6820	TW01-37	0.0066	Manual	T
679470.3570	3082776.7350	TW01-38	0.0105	Manual	T
679497.3310	3082840.3960	TW01-39	0.0027	Manual	T
679524.3310	3082886.8990	TW01-40	0.0027	Manual	T
679560.6110	3082897.2580	TW01-41	0.0075	Manual	T
679431.2599	3082589.5638		0	Random	

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680310.3290	3082668.1710	J-59SD	0.0027	Manual	T
680352.2560	3082820.3630	J-60SD	0.0027	Manual	T
680379.4090	3082937.1350	J-61SD	0.0027	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679386.3850	3083044.5490	TW06-63	0.0031	Manual	T
679396.8510	3083038.0640	TW06-64	0.0027	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**

Area: Area 5					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679133.4290	3083306.3130	TW01-01	0.0027	Manual	T
679104.2450	3083223.2620	TW01-02	0.0028	Manual	T
679242.7260	3083326.5280	TW01-07	0.0027	Manual	T
679181.2750	3083178.2880	TW01-08	0.0038	Manual	T
679268.7700	3083200.3260	TW01-11	0.0035	Manual	T
679301.1600	3083254.0340	TW01-12	0.0027	Manual	T

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability (1-β) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

- *n* is the number of samples,
- *S* is the estimated standard deviation of the measured values including analytical error,
- Δ is the width of the gray region,
- α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
- β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
- $Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is 1-α,
- $Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is 1-β.

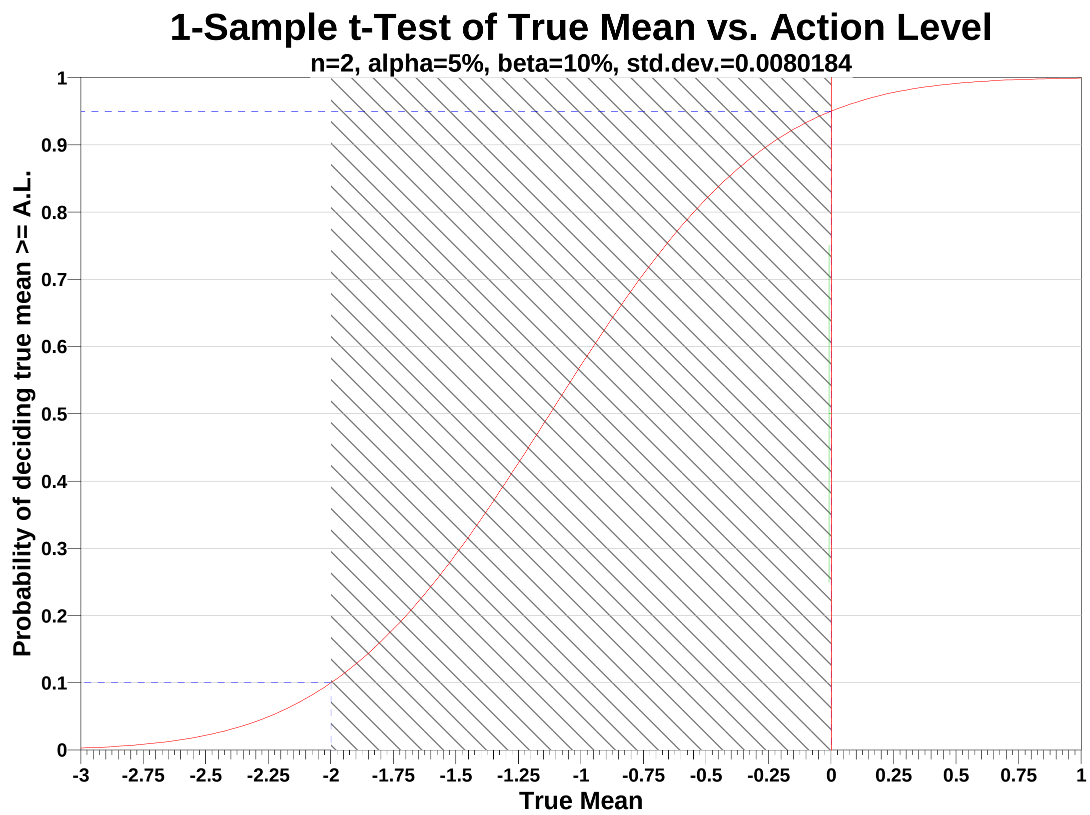
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	2	0.0080184	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=4.5e-005		α=5		α=10		α=15	
		s=0.0160368	s=0.0080184	s=0.0160368	s=0.0080184	s=0.0160368	s=0.0080184
LBGR=90	β=5	137443723	34360932	108762535	27190635	91305516	22826380
	β=10	108762535	27190635	83433904	20858477	68238839	17059711
	β=15	91305517	22826381	68238840	17059711	54569904	13642477
LBGR=80	β=5	34360932	8590234	27190635	6797660	22826380	5706596
	β=10	27190635	6797660	20858477	5214620	17059711	4264928
	β=15	22826381	5706597	17059711	4264929	13642477	3410620
LBGR=70	β=5	15271526	3817883	12084727	3021183	10145058	2536265

	β=10	12084728	3021183	9270435	2317610	7582094	1895524
	β=15	10145059	2536266	7582094	1895525	6063324	1515832

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $2,000.00, which averages out to a per sample cost of $1,000.00. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**2 Samples**
Field collection costs		$100.00	$200.00
Analytical costs	$400.00	$400.00	$800.00
Sum of Field & Analytical costs		**$500.00**	**$1,000.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$2,000.00**

Data Analysis
The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10
0	0	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027
10	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0028	0.0029
20	0.0031	0.0035	0.0038	0.0046	0.0066	0.007	0.0075	0.0076	0.0082	0.0085
30	0.0095	0.0105	0.0106	0.01665	0.0195	0.0211	0.0437			

SUMMARY STATISTICS	
n	37
Min	0
Max	0.0437
Range	0.0437
Mean	0.0065824
Median	0.0028
Variance	6.2938e-005
StdDev	0.0079334
Std Error	0.0013042
Skewness	3.2964
Interquartile Range	0.0052
Percentiles	

1%	5%	10%	25%	50%	75%	90%	95%	99%
0	0.00243	0.0027	0.0027	0.0028	0.0079	0.01722	0.02336	0.0437

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	4.616	2.99	Yes

The test statistic 4.616 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	0.0437

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Shapiro-Wilk test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Shapiro-Wilk Test Statistic	0.7187
Shapiro-Wilk 5% Critical Value	0.934

The calculated Shapiro-Wilk test statistic is less than the 5% Shapiro-Wilk critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

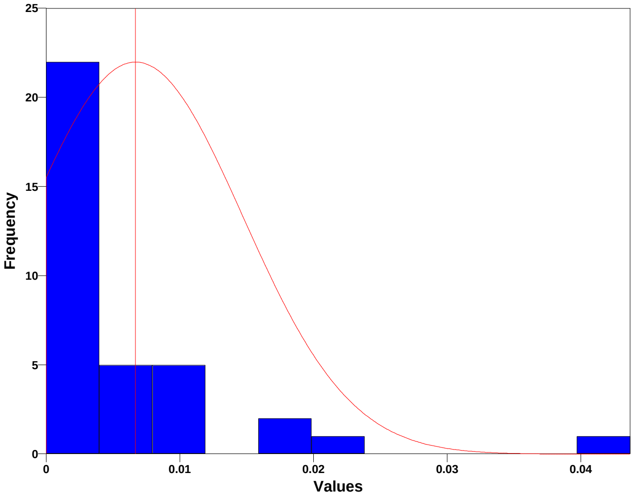
Data Plots
Graphical displays of the data are shown below.

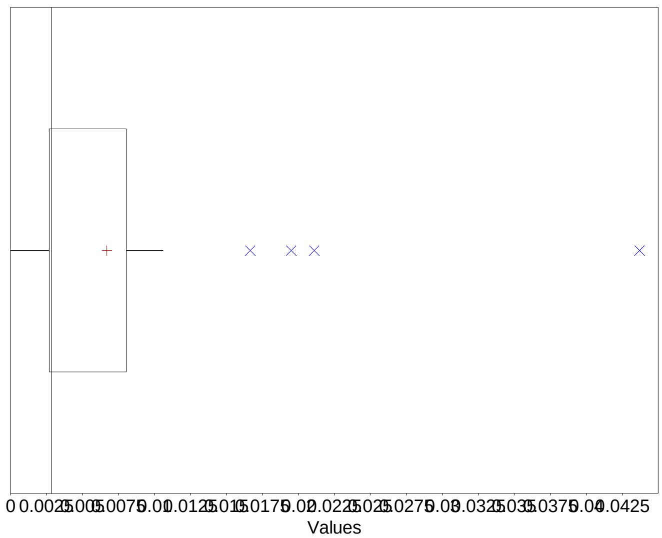
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

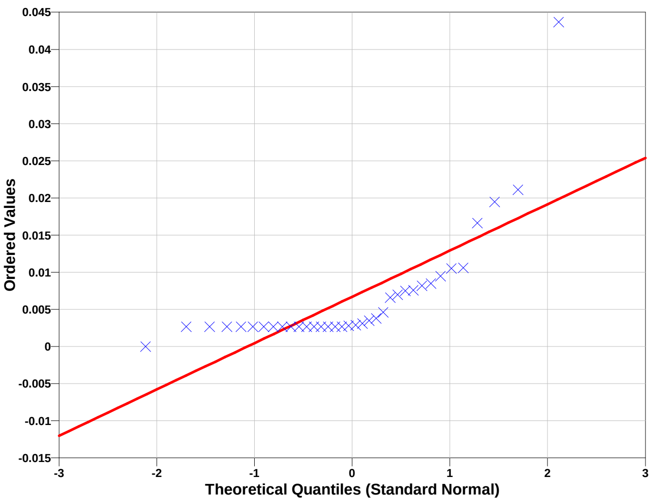
The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted

individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_n, for which a fraction p of the distribution is less than x_n. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Shapiro-Wilk (SW) test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Shapiro-Wilk Test Statistic	0.5979
Shapiro-Wilk 5% Critical Value	0.936

The calculated SW test statistic is less than the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance.

The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.008784
95% Non-Parametric (Chebyshev) UCL	0.01227

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.01227) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=37 data,
AL is the action level or threshold (4.5e-005),
SE is the standard error = (standard deviation) / (square root of n).

This *t* was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=36 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
5.0125	1.6883	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
1	23	Cannot Reject
Note: There may not be enough data to reject the null hypothesis (and conclude site is clean) with 95% confidence using the MARSSIM sign test.		

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

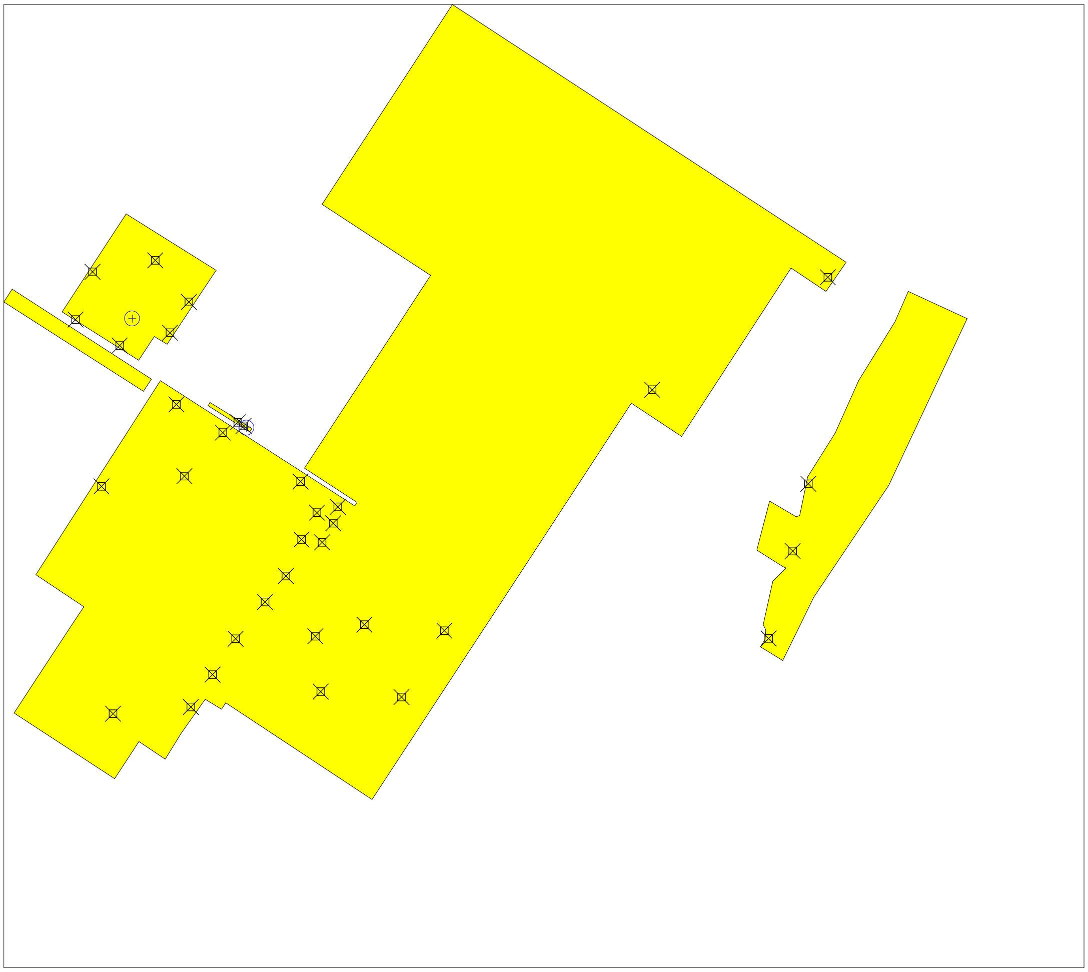
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	36
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679530.9430	3082575.4480	G-36SD	0.007	Manual	T
679606.6840	3082692.3830	G-37SD	0.0029	Manual	T
679671.3170	3082565.9250	G-46SD	0.0027	Manual	T
679745.9820	3082681.3860	G-47SD	0.0027	Manual	T
679532.9930	3082835.5820	J-42SD	0.0085	Manual	T
679552.9590	3082868.6600	J-43SD	0.0076	Manual	T
679521.7790	3082672.0220	J-54SD	0.0027	Manual	T
680108.0130	3083101.3520	J-57SD	0.0106	Manual	T
680413.8310	3083297.0130	J-58SD	0.0027	Manual	T
679149.4920	3082933.0980	TW01-13	0.0027	Manual	T
679279.7760	3083075.6320	TW01-14	0.0027	Manual	T
679293.5600	3082950.4980	TW01-17	0.0046	Manual	T
679360.5700	3083026.4980	TW01-18	0.0211	Manual	T
679169.0760	3082537.3510	TW01-27	0.01665	Manual	T
679495.8840	3082940.9730	TW01-33	0.0095	Manual	T
679304.6530	3082548.6880	TW01-34	0.0082	Manual	T

679342.7410	3082605.3190	TW01-35	0.0195	Manual	T
679382.8900	3082667.5270	TW01-36	0.0437	Manual	T
679433.9450	3082731.6820	TW01-37	0.0066	Manual	T
679470.3570	3082776.7350	TW01-38	0.0105	Manual	T
679497.3310	3082840.3960	TW01-39	0.0027	Manual	T
679524.3310	3082886.8990	TW01-40	0.0027	Manual	T
679560.6110	3082897.2580	TW01-41	0.0075	Manual	T

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680310.3290	3082668.1710	J-59SD	0.0027	Manual	T
680352.2560	3082820.3630	J-60SD	0.0027	Manual	T
680379.4090	3082937.1350	J-61SD	0.0027	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679386.3850	3083044.5490	TW06-63	0.0031	Manual	T
679396.8510	3083038.0640	TW06-64	0.0027	Manual	T
679401.5543	3083034.8156	TW01-01	0.0027	Random	

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**

Area: Area 5					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679133.4290	3083306.3130	TW01-01	0.0027	Manual	T
679104.2450	3083223.2620	TW01-02	0.0028	Manual	T
679242.7260	3083326.5280	TW01-07	0.0027	Manual	T
679181.2750	3083178.2880	TW01-08	0.0038	Manual	T
679268.7700	3083200.3260	TW01-11	0.0035	Manual	T
679301.1600	3083254.0340	TW01-12	0.0027	Manual	T
679202.5111	3083224.7611		0	Random	

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric

approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

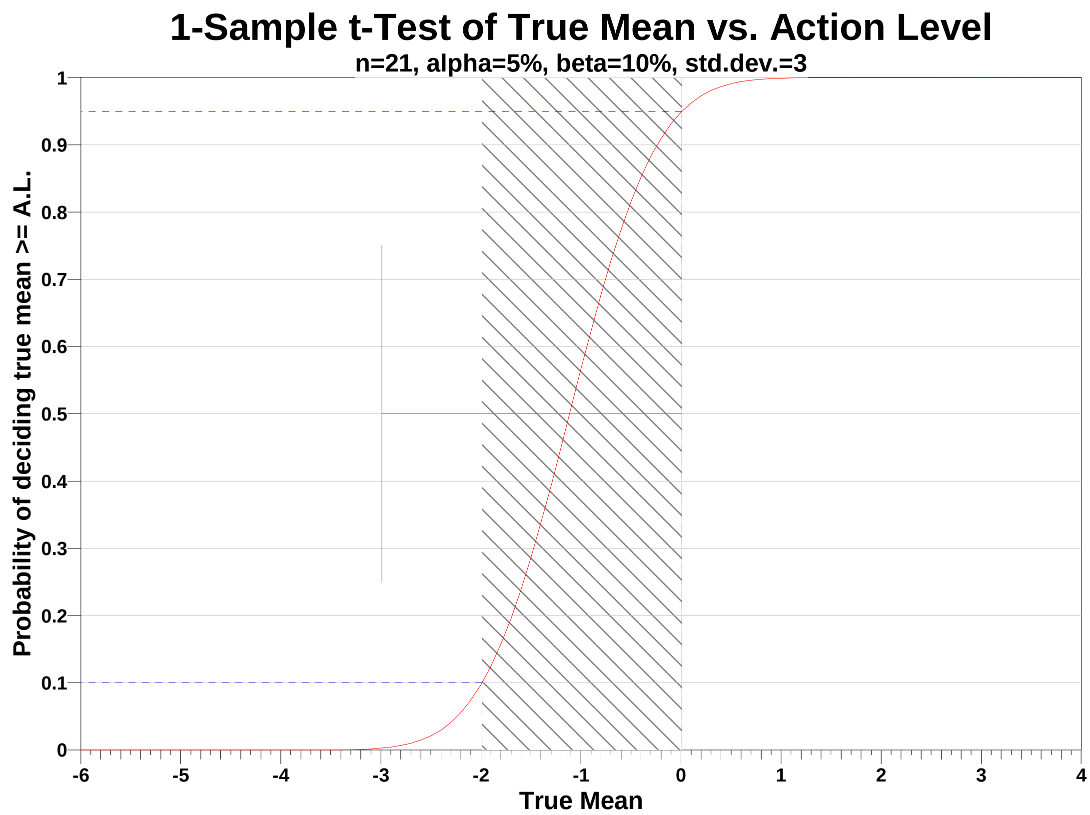
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.01		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	389598259	97399566	308298506	77074627	258814803	64703702
	β=10	308298506	77074628	236501917	59125480	193429954	48357489
	β=15	258814804	64703702	193429954	48357490	154683962	38670991
LBGR=80	**β=5**	97399566	24349893	77074627	19268658	64703702	16175926
	β=10	77074628	19268658	59125480	14781371	48357489	12089373
	β=15	64703702	16175927	48357490	12089373	38670991	9667749
LBGR=70	**β=5**	43288697	10822176	34255391	8563849	28757201	7189301

	β=10	34255391	8563849	26277992	6569499	21492218	5373055
	β=15	28757202	7189302	21492218	5373056	17187108	4296778

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis
The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10
0	0	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027
10	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0028	0.0029
20	0.0031	0.0031	0.0035	0.0038	0.0046	0.0066	0.007	0.0075	0.0076	0.0082
30	0.0085	0.0095	0.0105	0.0106	0.01665	0.0195	0.0211	0.0437		

SUMMARY STATISTICS	
n	38
Min	0
Max	0.0437
Range	0.0437
Mean	0.0064908
Median	0.00285
Variance	6.1556e-005
StdDev	0.0078458
Std Error	0.0012728
Skewness	3.3454
Interquartile Range	0.00505
Percentiles	

1%	5%	10%	25%	50%	75%	90%	95%	99%
0	0.002565	0.0027	0.0027	0.00285	0.00775	0.01694	0.02223	0.0437

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	4.68	3	Yes

The test statistic 4.68 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	0.0437

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Shapiro-Wilk test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Shapiro-Wilk Test Statistic	0.7114
Shapiro-Wilk 5% Critical Value	0.935

The calculated Shapiro-Wilk test statistic is less than the 5% Shapiro-Wilk critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

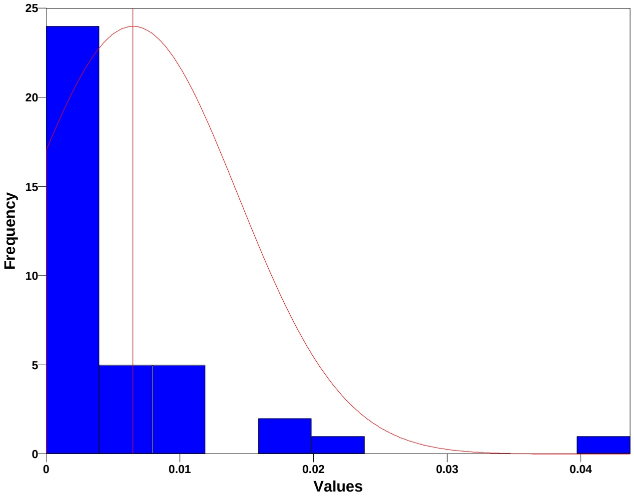
Data Plots
Graphical displays of the data are shown below.

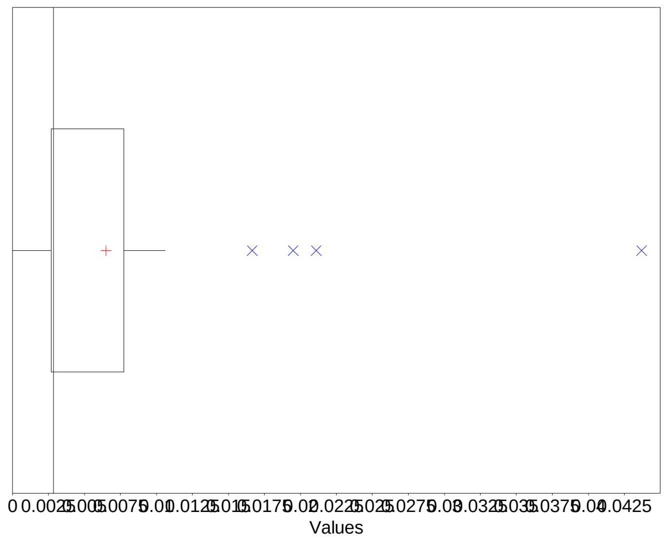
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

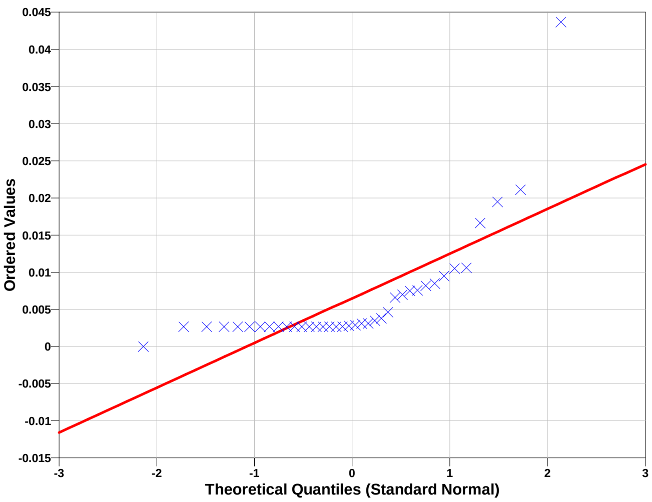
The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted

individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_n, for which a fraction p of the distribution is less than x_n. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Shapiro-Wilk (SW) test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Shapiro-Wilk Test Statistic	0.5919
Shapiro-Wilk 5% Critical Value	0.938

The calculated SW test statistic is less than the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance.

The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.008638
95% Non-Parametric (Chebyshev) UCL	0.01204

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.01204) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=38 data,
AL is the action level or threshold (0.01),
SE is the standard error = (standard deviation) / (square root of n).

This *t* was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=37 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-2.7572	1.6871	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
32	24	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

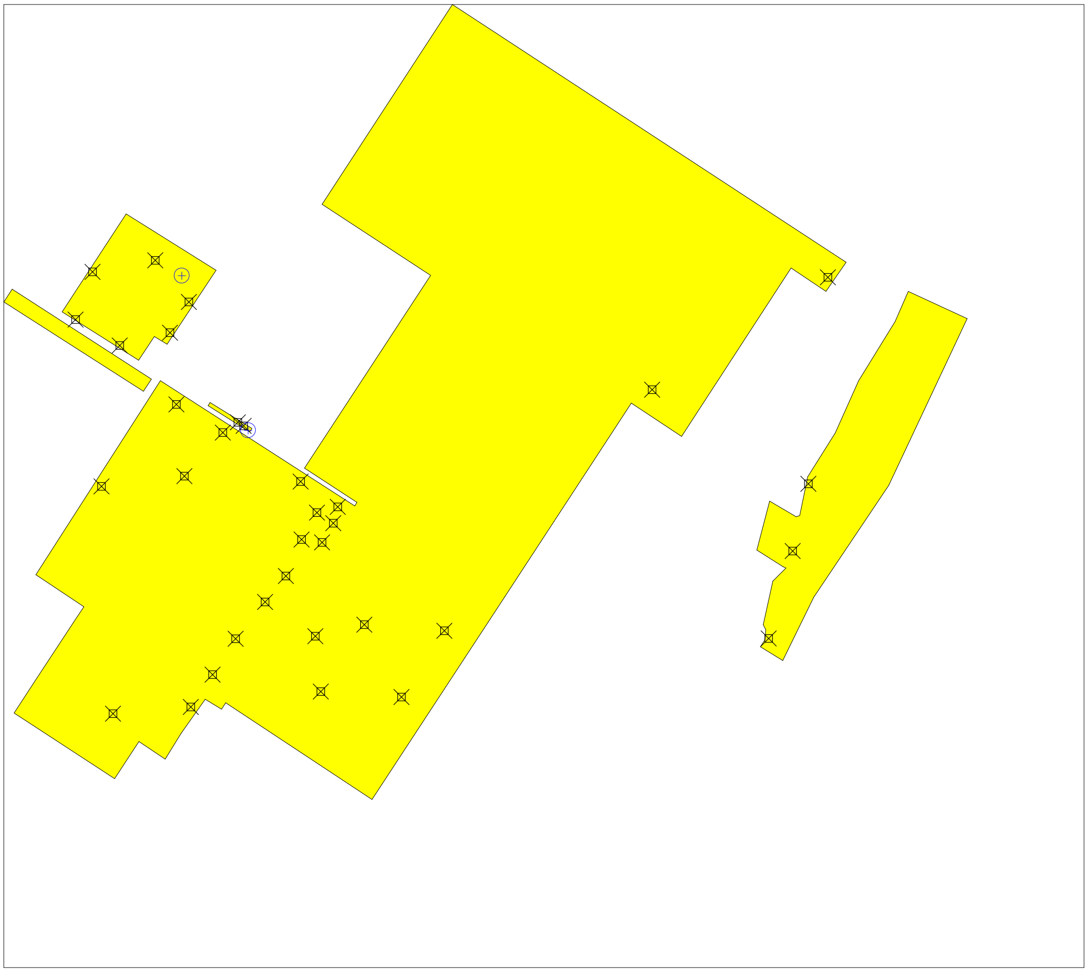
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	36
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	Y Coord	Label	Value	Type	Historical
679530.9430	3082575.4480	G-36SD	7	Manual	T
679606.6840	3082692.3830	G-37SD	2.9	Manual	T
679671.3170	3082565.9250	G-46SD	1.35	Manual	T
679745.9820	3082681.3860	G-47SD	1.35	Manual	T
679532.9930	3082835.5820	J-42SD	8.5	Manual	T
679552.9590	3082868.6600	J-43SD	7.6	Manual	T
679521.7790	3082672.0220	J-54SD	1.35	Manual	T
680108.0130	3083101.3520	J-57SD	10.6	Manual	T
680413.8310	3083297.0130	J-58SD	1.35	Manual	T
679149.4920	3082933.0980	TW01-13	1.35	Manual	T
679279.7760	3083075.6320	TW01-14	2.7	Manual	T
679293.5600	3082950.4980	TW01-17	4.6	Manual	T
679360.5700	3083026.4980	TW01-18	21.1	Manual	T
679169.0760	3082537.3510	TW01-27	16.65	Manual	T
679495.8840	3082940.9730	TW01-33	9.5	Manual	T
679304.6530	3082548.6880	TW01-34	8.2	Manual	T

679342.7410	3082605.3190	TW01-35	19.5	Manual	T
679382.8900	3082667.5270	TW01-36	43.7	Manual	T
679433.9450	3082731.6820	TW01-37	6.6	Manual	T
679470.3570	3082776.7350	TW01-38	10.5	Manual	T
679497.3310	3082840.3960	TW01-39	1.35	Manual	T
679524.3310	3082886.8990	TW01-40	1.35	Manual	T
679560.6110	3082897.2580	TW01-41	7.5	Manual	T

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680310.3290	3082668.1710	J-59SD	1.35	Manual	T
680352.2560	3082820.3630	J-60SD	1.35	Manual	T
680379.4090	3082937.1350	J-61SD	1.35	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679386.3850	3083044.5490	TW06-63	3.1	Manual	T
679396.8510	3083038.0640	TW06-64	1.35	Manual	T
679404.4014	3083030.3745	TW01-01	1.35	Random	

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**

Area: Area 5					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679133.4290	3083306.3130	TW01-01	1.35	Manual	T
679104.2450	3083223.2620	TW01-02	2.8	Manual	T
679242.7260	3083326.5280	TW01-07	1.35	Manual	T
679181.2750	3083178.2880	TW01-08	3.8	Manual	T
679268.7700	3083200.3260	TW01-11	3.5	Manual	T
679301.1600	3083254.0340	TW01-12	1.35	Manual	T
679289.1671	3083299.2675		0	Random	

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric

approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs

The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5 Z_{1-\alpha}^2$$

where

- n is the number of samples,
- S is the estimated standard deviation of the measured values including analytical error,
- Δ is the width of the gray region,
- α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
- β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
- $Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
- $Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

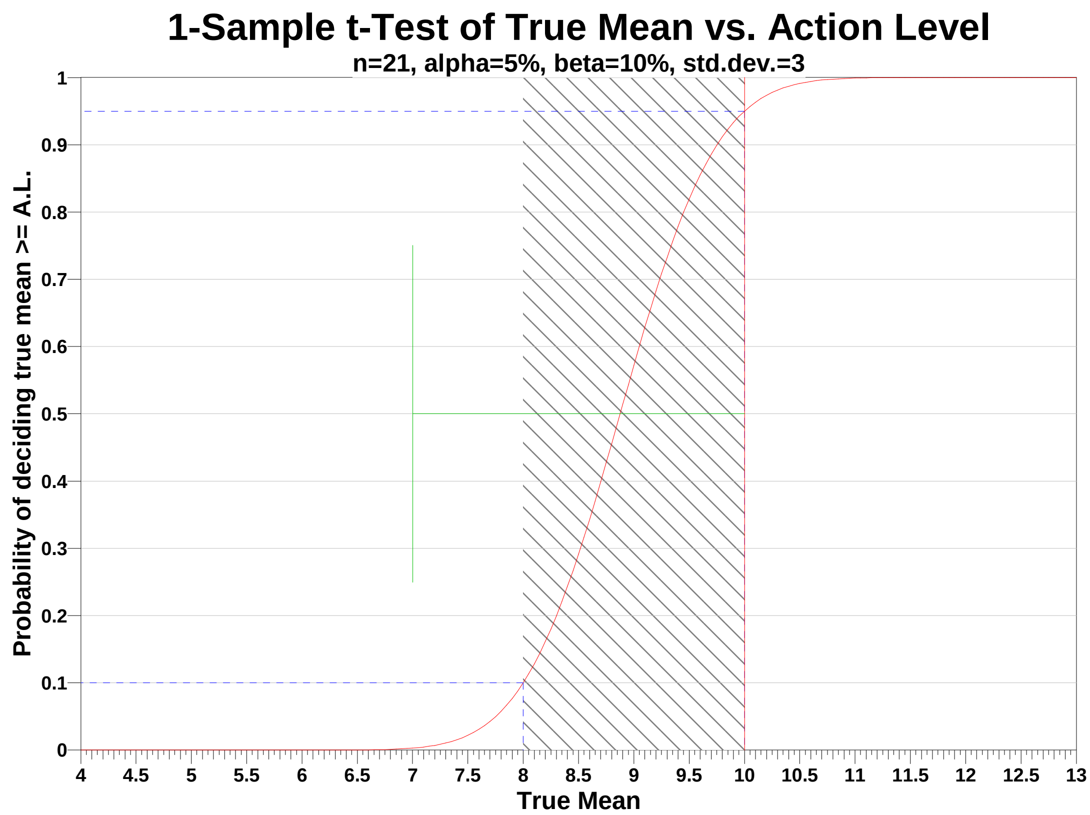
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=10		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	391	99	310	78	260	66
	β=10	310	79	238	60	194	49
	β=15	261	67	195	50	156	40
LBGR=80	β=5	99	26	78	21	66	17
	β=10	79	21	60	16	49	13
	β=15	67	18	50	13	40	11
LBGR=70	β=5	45	13	36	10	30	8

	β=10	36	10	28	8	23	6
	β=15	31	9	23	7	18	5

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that $\mu >$ action level
α = Alpha (%), Probability of mistakenly concluding that $\mu <$ action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis
The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10
0	0	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.35
10	1.35	1.35	1.35	1.35	1.35	1.35	2.7	2.8	2.9	3.1
20	3.5	3.8	4.6	6.6	7	7.5	7.6	8.2	8.5	9.5
30	10.5	10.6	16.65	19.5	21.1	43.7				

SUMMARY STATISTICS	
n	36
Min	0
Max	43.7
Range	43.7
Mean	6.1278
Median	2.85
Variance	69.367
StdDev	8.3287
Std Error	1.3881
Skewness	3.0451
Interquartile Range	6.7
Percentiles	

1%	5%	10%	25%	50%	75%	90%	95%	99%
0	1.148	1.35	1.35	2.85	8.05	17.5	24.49	43.7

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	4.401	2.97	Yes

The test statistic 4.401 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	43.7

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Shapiro-Wilk test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Shapiro-Wilk Test Statistic	0.7543
Shapiro-Wilk 5% Critical Value	0.931

The calculated Shapiro-Wilk test statistic is less than the 5% Shapiro-Wilk critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

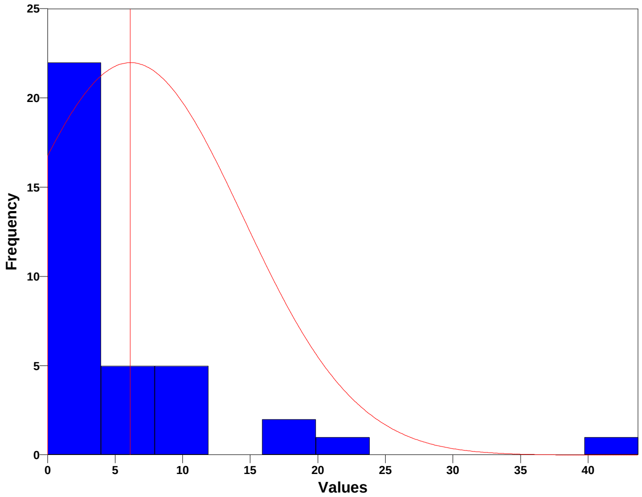
Data Plots
Graphical displays of the data are shown below.

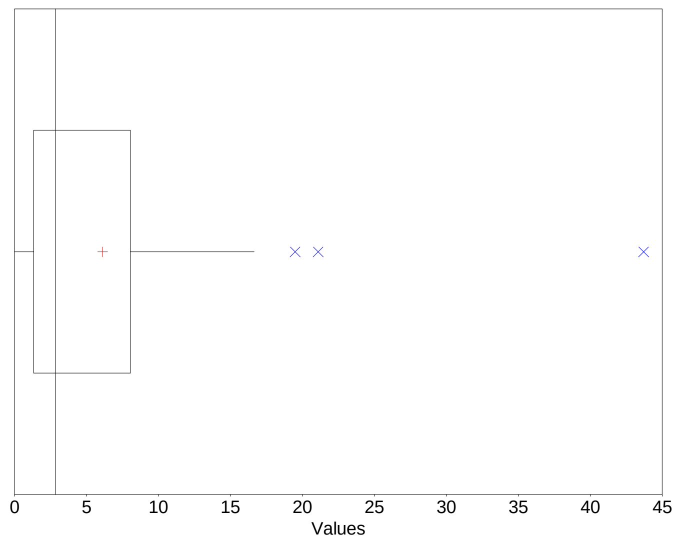
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

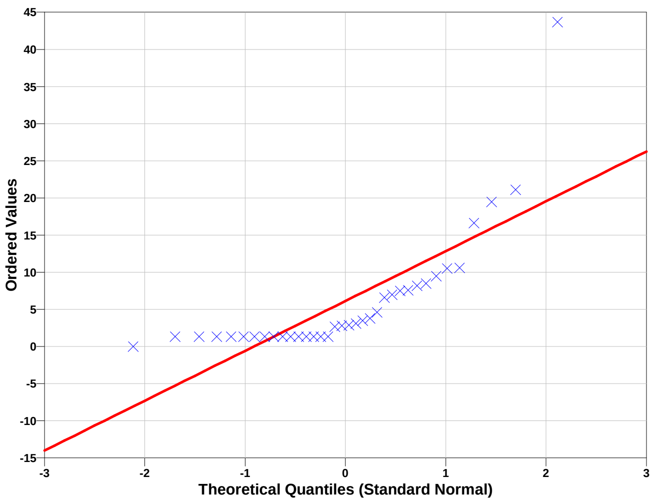
The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted

individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_n, for which a fraction p of the distribution is less than x_n. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Shapiro-Wilk (SW) test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Shapiro-Wilk Test Statistic	0.6436
Shapiro-Wilk 5% Critical Value	0.935

The calculated SW test statistic is less than the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance.

The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	8.473
95% Non-Parametric (Chebyshev) UCL	12.18

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (12.18) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=36 data,
AL is the action level or threshold (10),
SE is the standard error = (standard deviation) / (square root of n).

This *t* was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=35 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-2.7896	1.6896	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
30	23	Reject

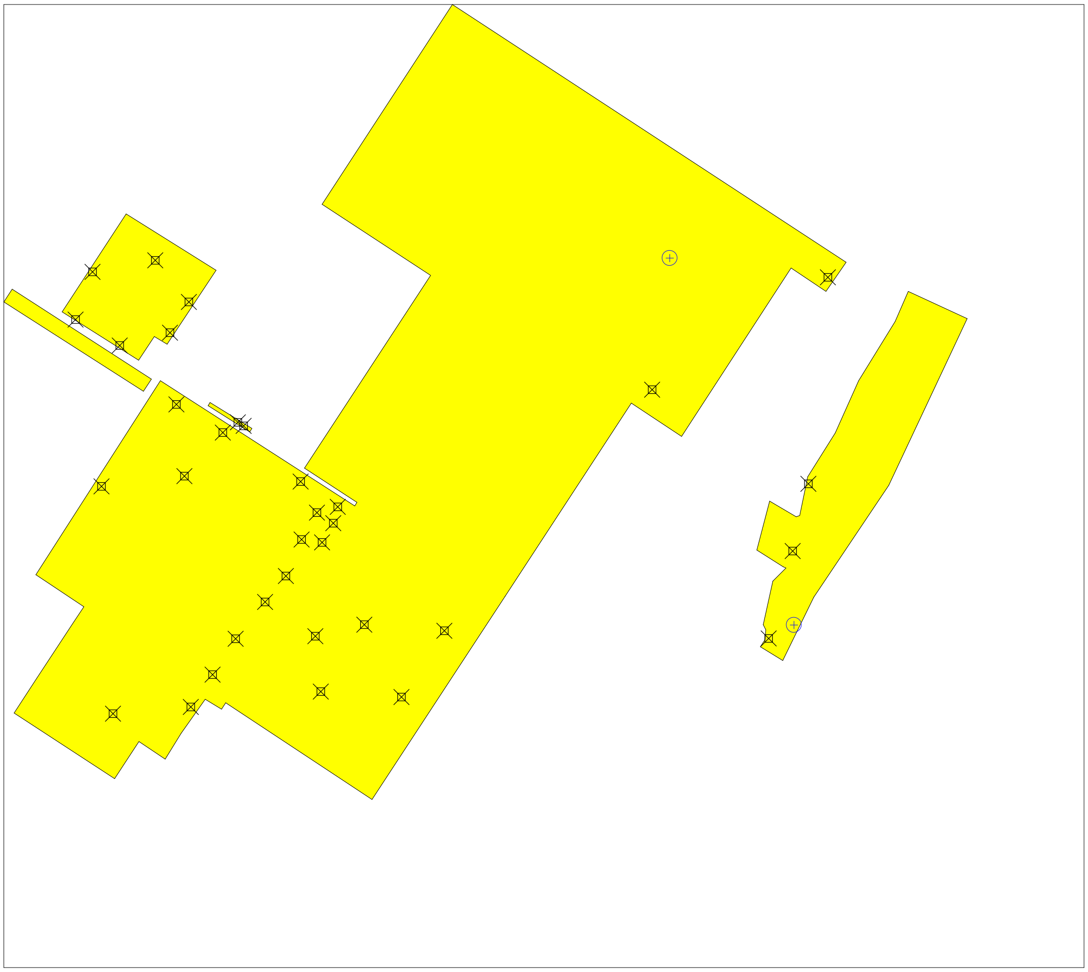
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	36
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	Y Coord	Label	Value	Type	Historical
679530.9430	3082575.4480	G-36SD	361000	Manual	T
679606.6840	3082692.3830	G-37SD	222000	Manual	T
679671.3170	3082565.9250	G-46SD	238000	Manual	T
679745.9820	3082681.3860	G-47SD	235000	Manual	T
679532.9930	3082835.5820	J-42SD	192000	Manual	T
679552.9590	3082868.6600	J-43SD	216000	Manual	T
679521.7790	3082672.0220	J-54SD	242500	Manual	T
680108.0130	3083101.3520	J-57SD	417000	Manual	T
680413.8310	3083297.0130	J-58SD	882000	Manual	T
679149.4920	3082933.0980	TW01-13	42700	Manual	T
679279.7760	3083075.6320	TW01-14	13300	Manual	T
679293.5600	3082950.4980	TW01-17	26100	Manual	T
679360.5700	3083026.4980	TW01-18	26500	Manual	T
679169.0760	3082537.3510	TW01-27	54050	Manual	T
679495.8840	3082940.9730	TW01-33	58100	Manual	T
679304.6530	3082548.6880	TW01-34	1.45e+006	Manual	T

679342.7410	3082605.3190	TW01-35	124000	Manual	T
679382.8900	3082667.5270	TW01-36	130000	Manual	T
679433.9450	3082731.6820	TW01-37	603000	Manual	T
679470.3570	3082776.7350	TW01-38	343000	Manual	T
679497.3310	3082840.3960	TW01-39	304000	Manual	T
679524.3310	3082886.8990	TW01-40	285000	Manual	T
679560.6110	3082897.2580	TW01-41	216000	Manual	T
680138.4983	3083329.9265	J-59SD	1.2e+006	Random	

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680310.3290	3082668.1710	J-59SD	1.2e+006	Manual	T
680352.2560	3082820.3630	J-60SD	1.165e+006	Manual	T
680379.4090	3082937.1350	J-61SD	1.18e+006	Manual	T
680354.4110	3082691.5331	TW06-63	15800	Random	

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679386.3850	3083044.5490	TW06-63	15800	Manual	T
679396.8510	3083038.0640	TW06-64	21700	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**

Area: Area 5					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679133.4290	3083306.3130	TW01-01	3820	Manual	T
679104.2450	3083223.2620	TW01-02	6800	Manual	T
679242.7260	3083326.5280	TW01-07	7770	Manual	T
679181.2750	3083178.2880	TW01-08	23600	Manual	T
679268.7700	3083200.3260	TW01-11	10300	Manual	T
679301.1600	3083254.0340	TW01-12	8520	Manual	T

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric

approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs

The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

- n is the number of samples,
- S is the estimated standard deviation of the measured values including analytical error,
- Δ is the width of the gray region,
- α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
- β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
- $Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
- $Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

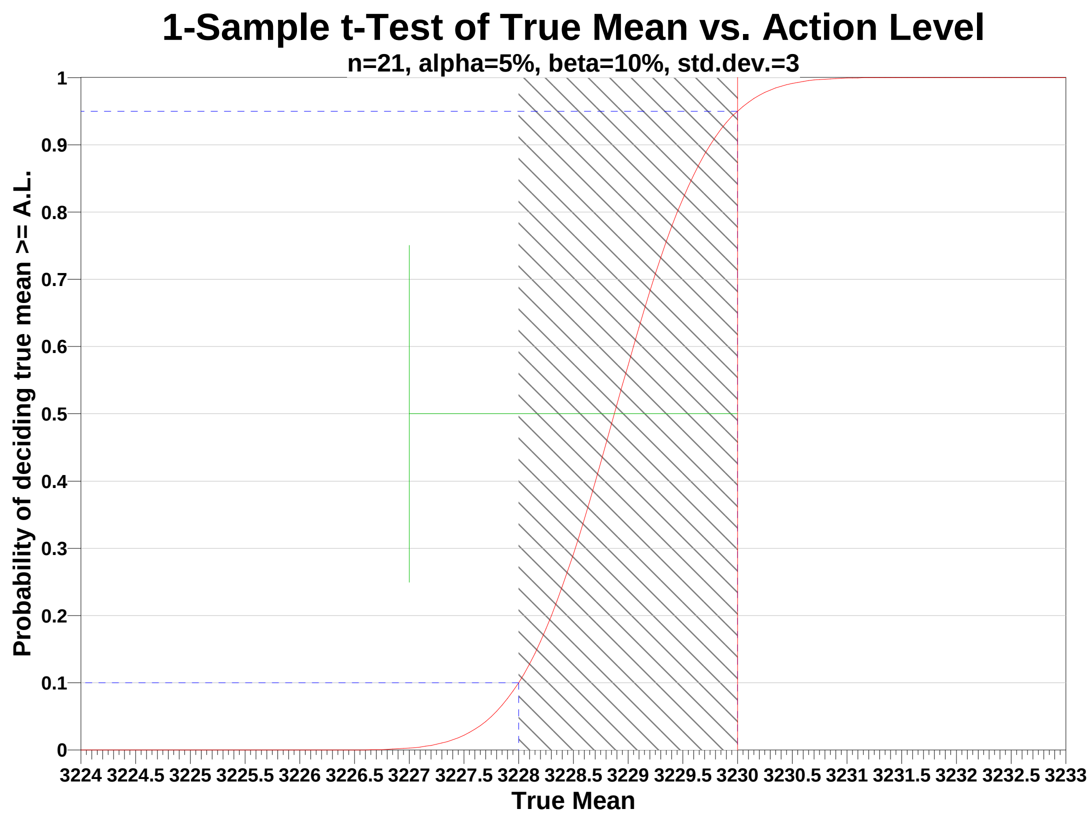
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=3230		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	2	2	1	1	1	1
	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1
LBGR=80	β=5	2	2	1	1	1	1
	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1
LBGR=70	β=5	2	2	1	1	1	1

	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis
The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10
0	3820	6800	7770	8520	1.03e+004	1.33e+004	1.58e+004	1.58e+004	2.17e+004	2.36e+004
10	2.61e+004	2.65e+004	4.27e+004	5.405e+004	5.81e+004	1.24e+005	1.3e+005	1.92e+005	2.16e+005	2.16e+005
20	2.22e+005	2.35e+005	2.38e+005	2.425e+005	2.85e+005	3.04e+005	3.43e+005	3.61e+005	4.17e+005	6.03e+005
30	8.82e+005	1.165e+006	1.18e+006	1.2e+006	1.2e+006	1.45e+006				

SUMMARY STATISTICS	
n	36
Min	3820
Max	1450000
Range	1.4462e+006
Mean	3.2057e+005
Median	204000
Variance	1.761e+011
StdDev	4.1965e+005
Std Error	69941
Skewness	1.5619
Interquartile Range	3.3433e+005
Percentiles	

1%	5%	10%	25%	50%	75%	90%	95%	99%
3820	6353	8295	2.218e+004	2.04e+005	3.565e+005	1.186e+006	1.237e+006	1.45e+006

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	2.864	2.97	No

None of the test statistics exceeded the corresponding critical values, therefore none of the 1 tests are significant and we conclude that at the 5% significance level there are no outliers in the data.

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Shapiro-Wilk test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Shapiro-Wilk Test Statistic	0.7187
Shapiro-Wilk 5% Critical Value	0.931

The calculated Shapiro-Wilk test statistic is less than the 5% Shapiro-Wilk critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

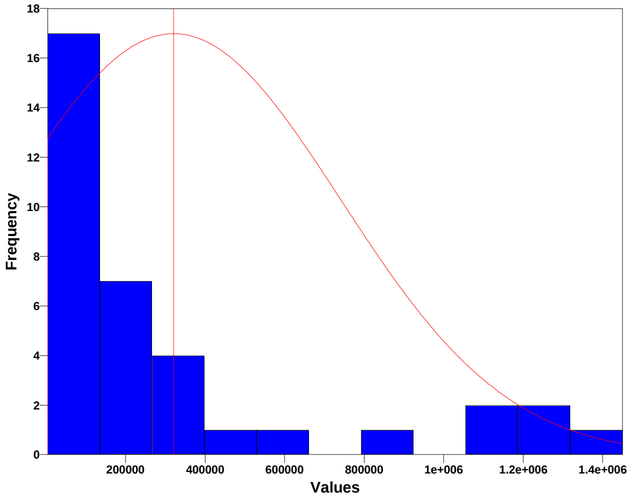
Data Plots
Graphical displays of the data are shown below.

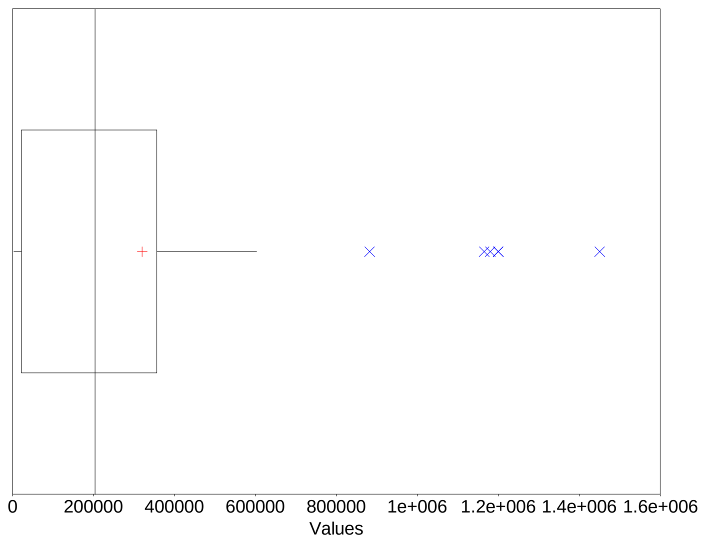
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

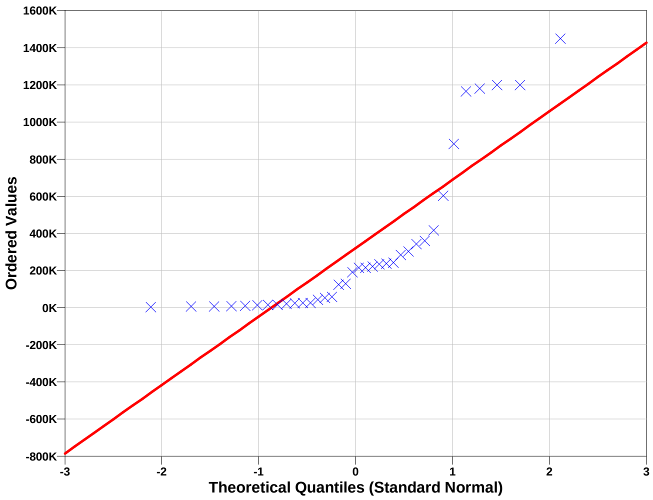
The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the

Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Shapiro-Wilk (SW) test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Shapiro-Wilk Test Statistic	0.727
Shapiro-Wilk 5% Critical Value	0.935

The calculated SW test statistic is less than the 5% Shapiro-Wilk critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	4.387e+005
95% Non-Parametric (Chebyshev) UCL	6.254e+005

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (6.254e+005) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=36 data,
AL is the action level or threshold (3230),
SE is the standard error = (standard deviation) / (square root of n).

This *t* was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=35 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
4.5372	1.6896	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
0	23	Cannot Reject
Note: There may not be enough data to reject the null hypothesis (and conclude site is clean) with 95% confidence using the MARSSIM sign test.		

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

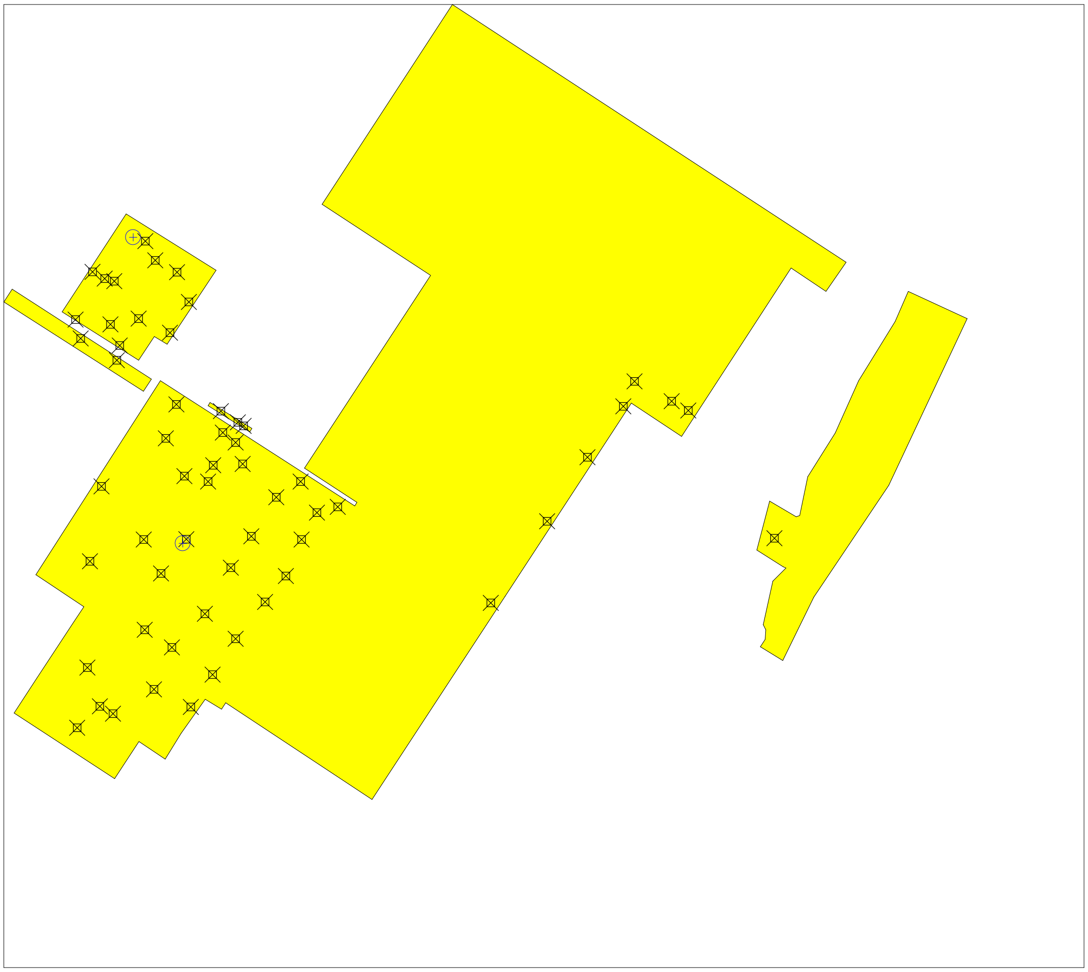
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679149.4920	3082933.0980	J-13S	0.54	Manual	T
679279.6830	3083075.4290	J-14S	0.23	Manual	T
679261.0980	3083016.3510	J-15S	0.38	Manual	T
679222.6340	3082840.1720	J-16S	0.33	Manual	T
679293.5600	3082950.4980	J-17S	0.41	Manual	T
679360.5700	3083026.4980	J-18S	0.44	Manual	T
679343.5810	3082969.5980	J-19S	1.7	Manual	T
679382.8640	3083009.1130	J-20S	2.4	Manual	T
679335.0020	3082941.1720	J-21S	0.2	Manual	T
679252.7130	3082781.0290	J-22S	0.74	Manual	T
679297.0010	3082840.6970	J-23S	1.4	Manual	T
679394.8070	3082971.8300	J-24S	0.3	Manual	T
679146.6460	3082549.7640	J-25S	1	Manual	T
679224.5850	3082683.1400	J-26S	0.22	Manual	T
679169.0760	3082537.3510	J-27S	1	Manual	T
679272.0040	3082652.6750	J-28S	1.2	Manual	T

679329.4380	3082711.0960	J-29S	0.58	Manual	T
679374.4420	3082791.3300	J-30S	0.54	Manual	T
679410.1490	3082845.8460	J-31S	0.23	Manual	T
679453.4760	3082914.1150	J-32S	1.1	Manual	T
679495.8840	3082940.9730	J-33S	0.81	Manual	T
679304.6530	3082548.6880	J-34S	0.8	Manual	T
679342.7410	3082605.3190	J-35S	0.78	Manual	T
679382.8900	3082667.5270	J-36S	1.7	Manual	T
679433.9450	3082731.6820	J-37S	2.1	Manual	T
679470.3570	3082776.7350	J-38S	2.2	Manual	T
679497.3310	3082840.3960	J-39S	0.7	Manual	T
679524.3310	3082886.8990	J-40S	0.62	Manual	T
679560.6070	3082897.2580	J-41S	1.7	Manual	T
679924.8150	3082872.3490	J-47S	1.3	Manual	T
679994.9690	3082983.5100	J-48S	1.2	Manual	T
680057.6580	3083072.0750	J-49S	0.69	Manual	T
680077.3540	3083115.5330	J-50S	0.52	Manual	T
679827.1150	3082729.7460	J-51S	0.63	Manual	T
680141.8730	3083080.8800	J-52S	1.1	Manual	T
680170.5600	3083064.6740	J-53S	2.4	Manual	T
679129.3320	3082802.5620	Composite 1	0.22	Manual	T
679124.7500	3082617.3010	Composite 3	1.6	Manual	T
679107.0750	3082512.5600	Composite 4	0.25	Manual	T
679240.6200	3082579.3320	Composite 2	0.44	Manual	T
679290.8532	3082833.4033	Composite 5	0.94	Random	

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.94	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.22	Manual	T
679396.8510	3083038.0640	J-64S	0.24	Manual	T
679386.3850	3083044.5490	J-63S	0.35	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.85	Manual	T
679113.1200	3083190.3150	J-66S	0.18	Manual	T

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.28	Manual	T
679104.2450	3083223.2620	J-02S	0.34	Manual	T
679155.0740	3083294.6960	J-03S	0.2	Manual	T
679171.2970	3083289.7960	J-04S	0.23	Manual	T
679225.8560	3083359.9740	J-05S	0.16	Manual	T
679164.8060	3083214.7100	J-06S	0.23	Manual	T
679242.7260	3083326.5280	J-07S	0.47	Manual	T
679181.2750	3083178.2880	J-08S	0.29	Manual	T
679213.7730	3083224.9730	J-09S	0.27	Manual	T
679280.5440	3083305.6810	J-10S	0.27	Manual	T
679268.7700	3083200.3260	J-11S	0.26	Manual	T
679301.1600	3083254.0340	J-12S	0.18	Manual	T
679204.6866	3083366.2710		0	Random	

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,

Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is 1-α,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is 1-β.

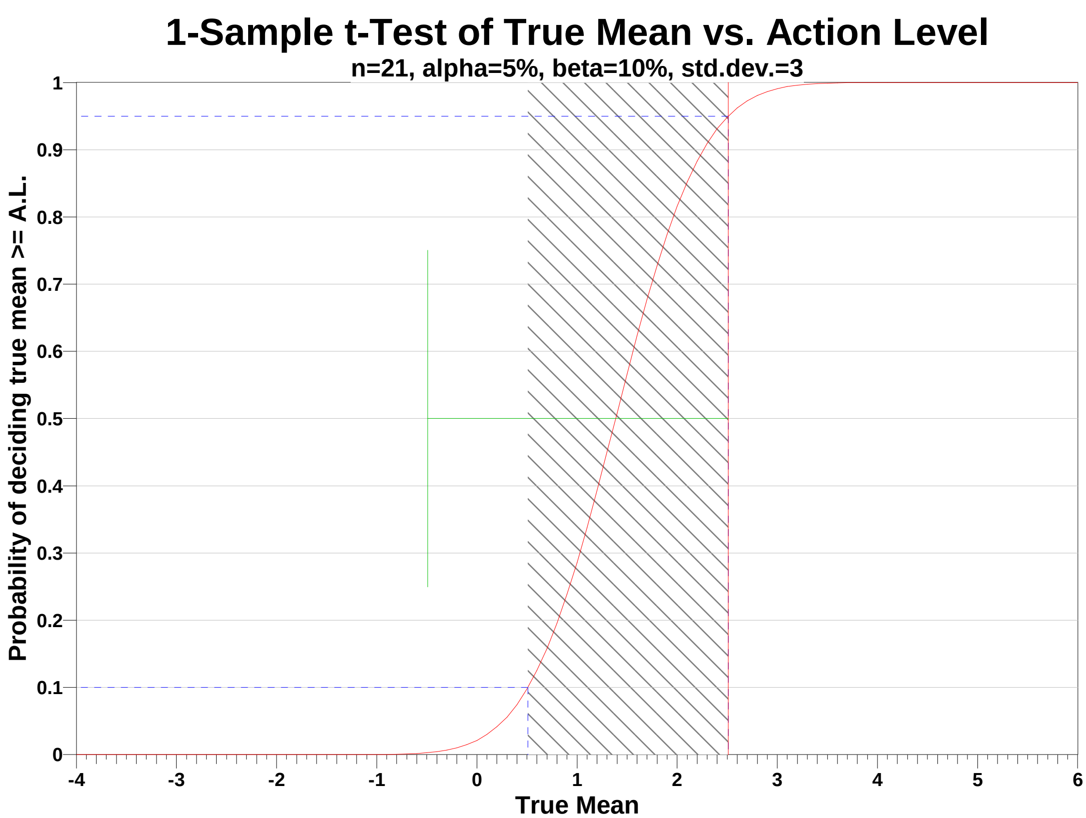
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30

or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=2.50958		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	6188	1548	4897	1225	4111	1028
	β=10	4897	1226	3757	940	3072	769
	β=15	4111	1029	3073	769	2457	615
LBGR=80	β=5	1548	388	1225	307	1028	258
	β=10	1226	308	940	236	769	193
	β=15	1029	259	769	193	615	155
LBGR=70	β=5	689	174	545	137	458	115
	β=10	546	138	419	106	342	86
	β=15	458	116	343	87	274	69

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling

The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10

0	0	0.16	0.18	0.18	0.2	0.2	0.22	0.22	0.22	0.22
10	0.23	0.23	0.23	0.23	0.24	0.25	0.26	0.27	0.27	0.28
20	0.28	0.29	0.3	0.33	0.34	0.35	0.38	0.41	0.44	0.44
30	0.47	0.52	0.54	0.54	0.58	0.62	0.63	0.69	0.7	0.74
40	0.78	0.8	0.81	0.85	0.85	0.94	0.94	1	1	1.1
50	1.1	1.2	1.2	1.3	1.4	1.6	1.7	1.7	1.7	2.1
60	2.2	2.4	2.4							

SUMMARY STATISTICS	
n	63
Min	0
Max	2.4
Range	2.4
Mean	0.71349
Median	0.52
Variance	0.35464
StdDev	0.59552
Std Error	0.075028
Skewness	1.3367
Interquartile Range	0.75

Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.18	0.208	0.25	0.52	1	1.7	2.18	2.4

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	2.832	3.218	No

None of the test statistics exceeded the corresponding critical values, therefore none of the 1 tests are significant and we conclude that at the 5% significance level there are no outliers in the data.

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)

Lilliefors Test Statistic	0.1573
Lilliefors 5% Critical Value	0.1125

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

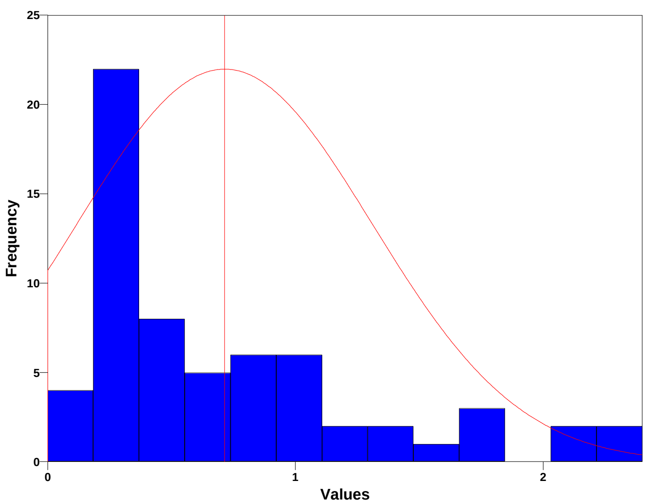
Data Plots

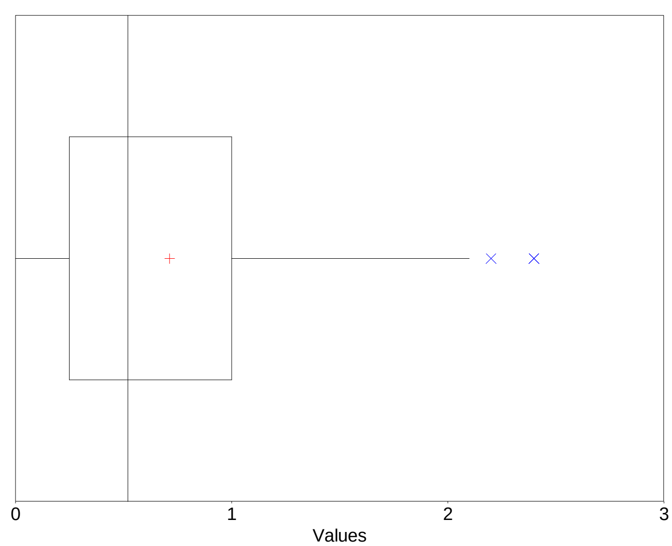
Graphical displays of the data are shown below.

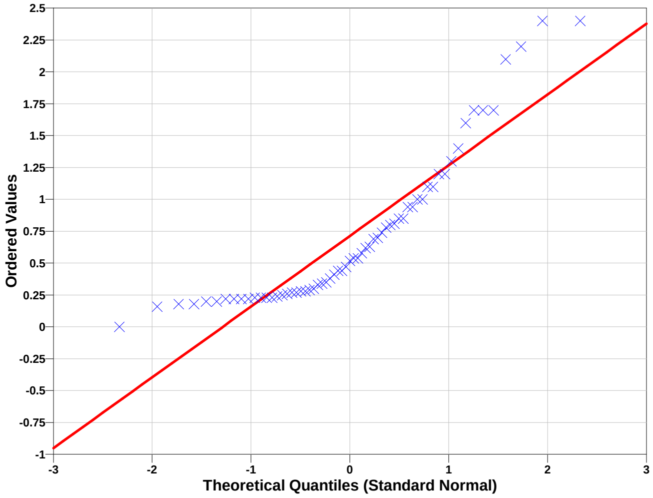
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.1605
Lilliefors 5% Critical Value	0.1116

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.8388
95% Non-Parametric (Chebyshev) UCL	1.041

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (1.041) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=63 data,
AL is the action level or threshold (2.50958),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=62 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-23.939	1.6698	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
63	39	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

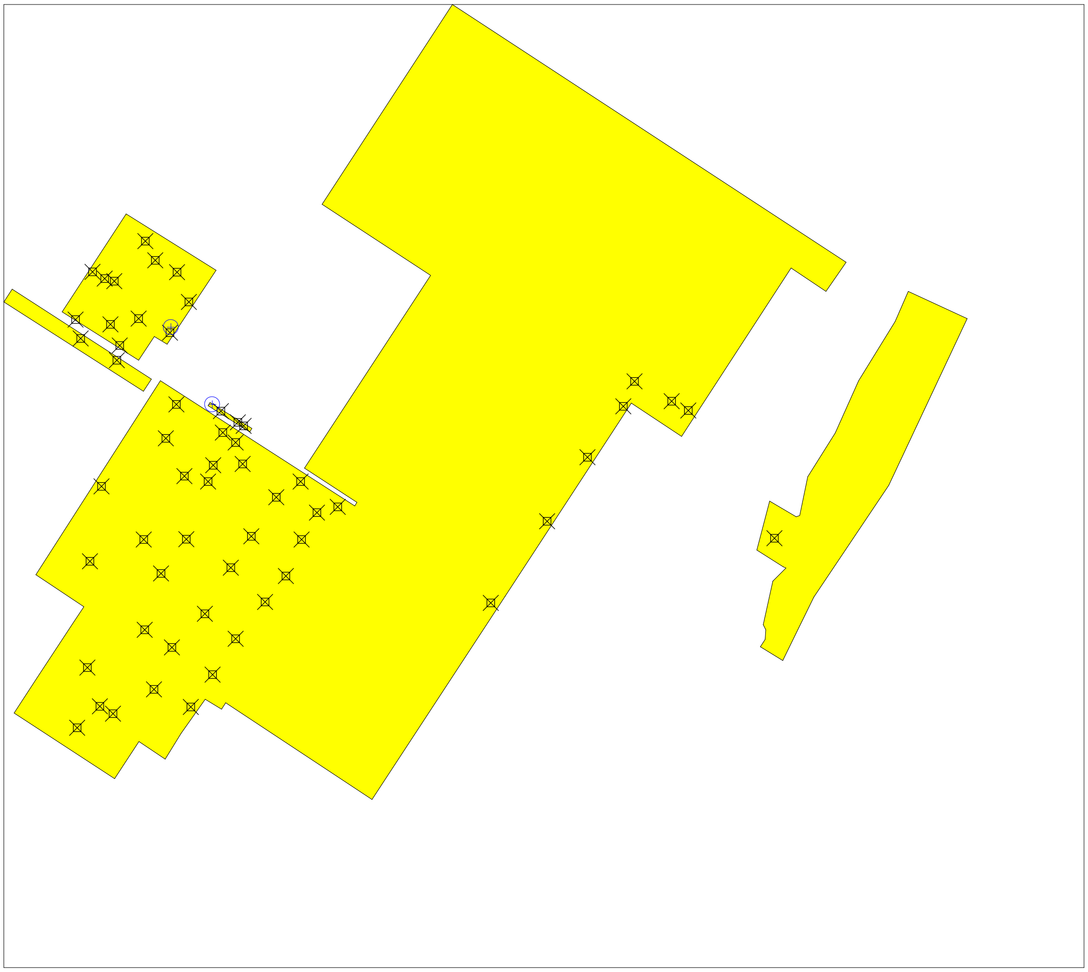
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679149.4920	3082933.0980	J-13S	0.54	Manual	T
679279.6830	3083075.4290	J-14S	0.23	Manual	T
679261.0980	3083016.3510	J-15S	0.38	Manual	T
679222.6340	3082840.1720	J-16S	0.33	Manual	T
679293.5600	3082950.4980	J-17S	0.41	Manual	T
679360.5700	3083026.4980	J-18S	0.44	Manual	T
679343.5810	3082969.5980	J-19S	1.7	Manual	T
679382.8640	3083009.1130	J-20S	2.4	Manual	T
679335.0020	3082941.1720	J-21S	0.2	Manual	T
679252.7130	3082781.0290	J-22S	0.74	Manual	T
679297.0010	3082840.6970	J-23S	1.4	Manual	T
679394.8070	3082971.8300	J-24S	0.3	Manual	T
679146.6460	3082549.7640	J-25S	1	Manual	T
679224.5850	3082683.1400	J-26S	0.22	Manual	T
679169.0760	3082537.3510	J-27S	1	Manual	T
679272.0040	3082652.6750	J-28S	1.2	Manual	T

679329.4380	3082711.0960	J-29S	0.58	Manual	T
679374.4420	3082791.3300	J-30S	0.54	Manual	T
679410.1490	3082845.8460	J-31S	0.23	Manual	T
679453.4760	3082914.1150	J-32S	1.1	Manual	T
679495.8840	3082940.9730	J-33S	0.81	Manual	T
679304.6530	3082548.6880	J-34S	0.8	Manual	T
679342.7410	3082605.3190	J-35S	0.78	Manual	T
679382.8900	3082667.5270	J-36S	1.7	Manual	T
679433.9450	3082731.6820	J-37S	2.1	Manual	T
679470.3570	3082776.7350	J-38S	2.2	Manual	T
679497.3310	3082840.3960	J-39S	0.7	Manual	T
679524.3310	3082886.8990	J-40S	0.62	Manual	T
679560.6070	3082897.2580	J-41S	1.7	Manual	T
679924.8150	3082872.3490	J-47S	1.3	Manual	T
679994.9690	3082983.5100	J-48S	1.2	Manual	T
680057.6580	3083072.0750	J-49S	0.69	Manual	T
680077.3540	3083115.5330	J-50S	0.52	Manual	T
679827.1150	3082729.7460	J-51S	0.63	Manual	T
680141.8730	3083080.8800	J-52S	1.1	Manual	T
680170.5600	3083064.6740	J-53S	2.4	Manual	T
679129.3320	3082802.5620	Composite 1	0.22	Manual	T
679124.7500	3082617.3010	Composite 3	1.6	Manual	T
679107.0750	3082512.5600	Composite 4	0.25	Manual	T
679240.6200	3082579.3320	Composite 2	0.44	Manual	T

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.94	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.22	Manual	T
679396.8510	3083038.0640	J-64S	0.24	Manual	T
679386.3850	3083044.5490	J-63S	0.35	Manual	T
679341.9051	3083075.2504	J-65S	0.85	Random	

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.85	Manual	T
679113.1200	3083190.3150	J-66S	0.18	Manual	T

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.28	Manual	T
679104.2450	3083223.2620	J-02S	0.34	Manual	T
679155.0740	3083294.6960	J-03S	0.2	Manual	T
679171.2970	3083289.7960	J-04S	0.23	Manual	T
679225.8560	3083359.9740	J-05S	0.16	Manual	T
679164.8060	3083214.7100	J-06S	0.23	Manual	T
679242.7260	3083326.5280	J-07S	0.47	Manual	T
679181.2750	3083178.2880	J-08S	0.29	Manual	T
679213.7730	3083224.9730	J-09S	0.27	Manual	T
679280.5440	3083305.6810	J-10S	0.27	Manual	T
679268.7700	3083200.3260	J-11S	0.26	Manual	T
679301.1600	3083254.0340	J-12S	0.18	Manual	T
679270.2667	3083209.5643		0	Random	

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,

Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

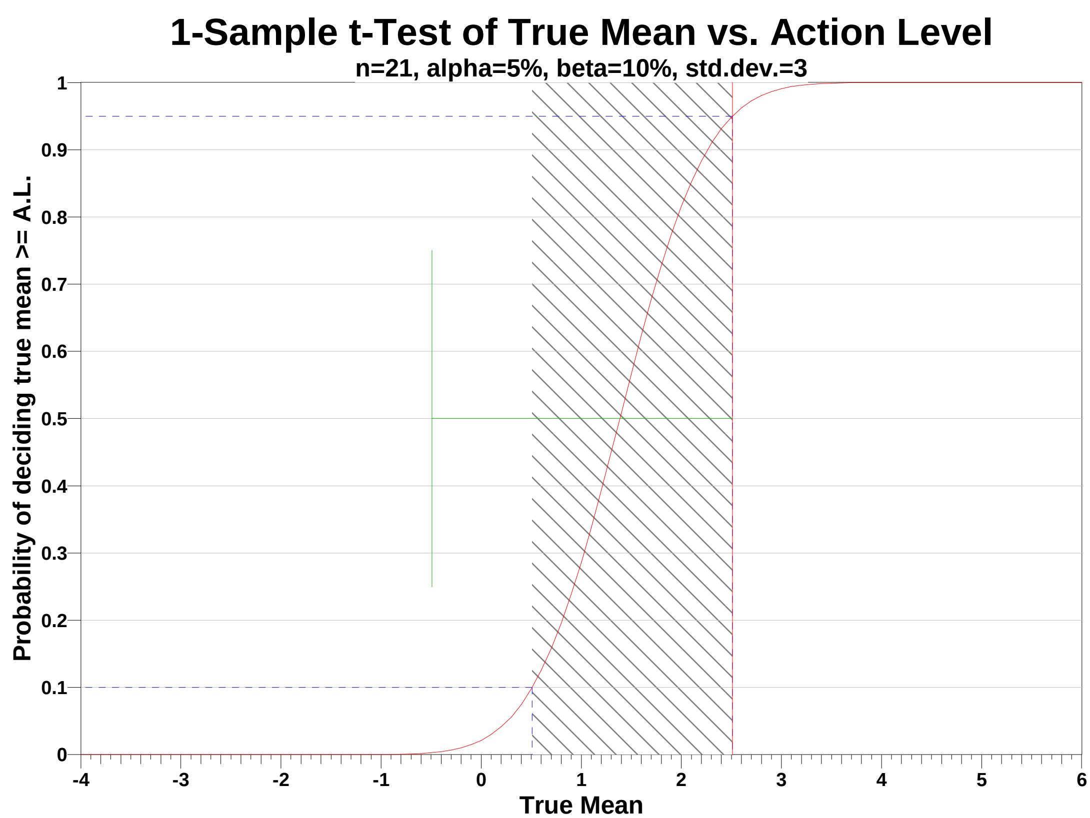
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30

or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=2.50958		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	6188	1548	4897	1225	4111	1028
	β=10	4897	1226	3757	940	3072	769
	β=15	4111	1029	3073	769	2457	615
LBGR=80	β=5	1548	388	1225	307	1028	258
	β=10	1226	308	940	236	769	193
	β=15	1029	259	769	193	615	155
LBGR=70	β=5	689	174	545	137	458	115
	β=10	546	138	419	106	342	86
	β=15	458	116	343	87	274	69

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling

The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

The following data points were entered by the user for analysis.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**

0	0	0.16	0.18	0.18	0.2	0.2	0.22	0.22	0.22	0.22
10	0.23	0.23	0.23	0.23	0.24	0.25	0.26	0.27	0.27	0.28
20	0.28	0.29	0.3	0.33	0.34	0.35	0.38	0.41	0.44	0.44
30	0.47	0.52	0.54	0.54	0.58	0.62	0.63	0.69	0.7	0.74
40	0.78	0.8	0.81	0.85	0.85	0.94	0.94	1	1	1.1
50	1.1	1.2	1.2	1.3	1.4	1.6	1.7	1.7	1.7	2.1
60	2.2	2.4	2.4							

SUMMARY STATISTICS	
n	63
Min	0
Max	2.4
Range	2.4
Mean	0.71349
Median	0.52
Variance	0.35464
StdDev	0.59552
Std Error	0.075028
Skewness	1.3367
Interquartile Range	0.75

Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.18	0.208	0.25	0.52	1	1.7	2.18	2.4

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	2.832	3.218	No

None of the test statistics exceeded the corresponding critical values, therefore none of the 1 tests are significant and we conclude that at the 5% significance level there are no outliers in the data.

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)

Lilliefors Test Statistic	0.1573
Lilliefors 5% Critical Value	0.1125

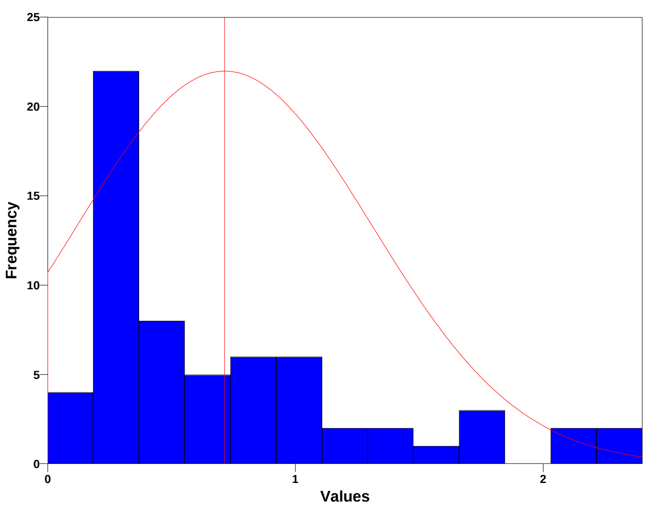
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

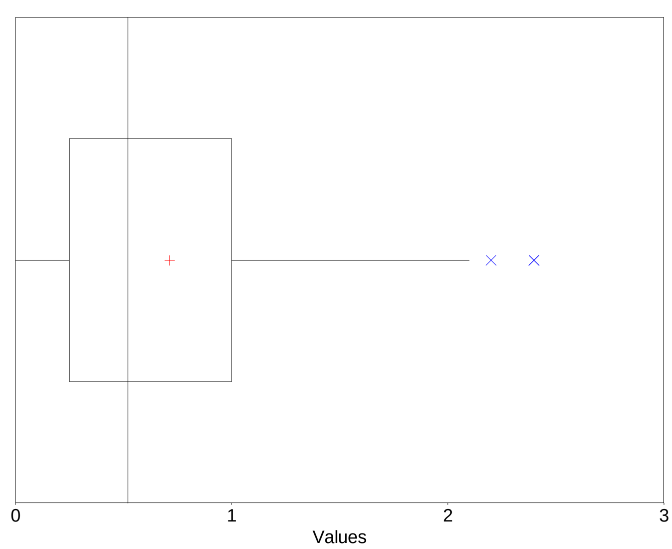
Data Plots
Graphical displays of the data are shown below.

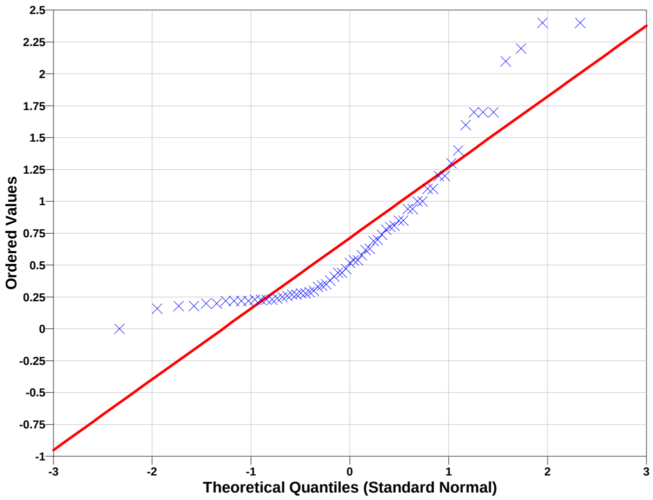
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25th and 75th percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The pth quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.1605
Lilliefors 5% Critical Value	0.1116

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.8388
95% Non-Parametric (Chebyshev) UCL	1.041

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (1.041) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=63 data,
AL is the action level or threshold (2.50958),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=62 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-23.939	1.6698	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
63	39	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

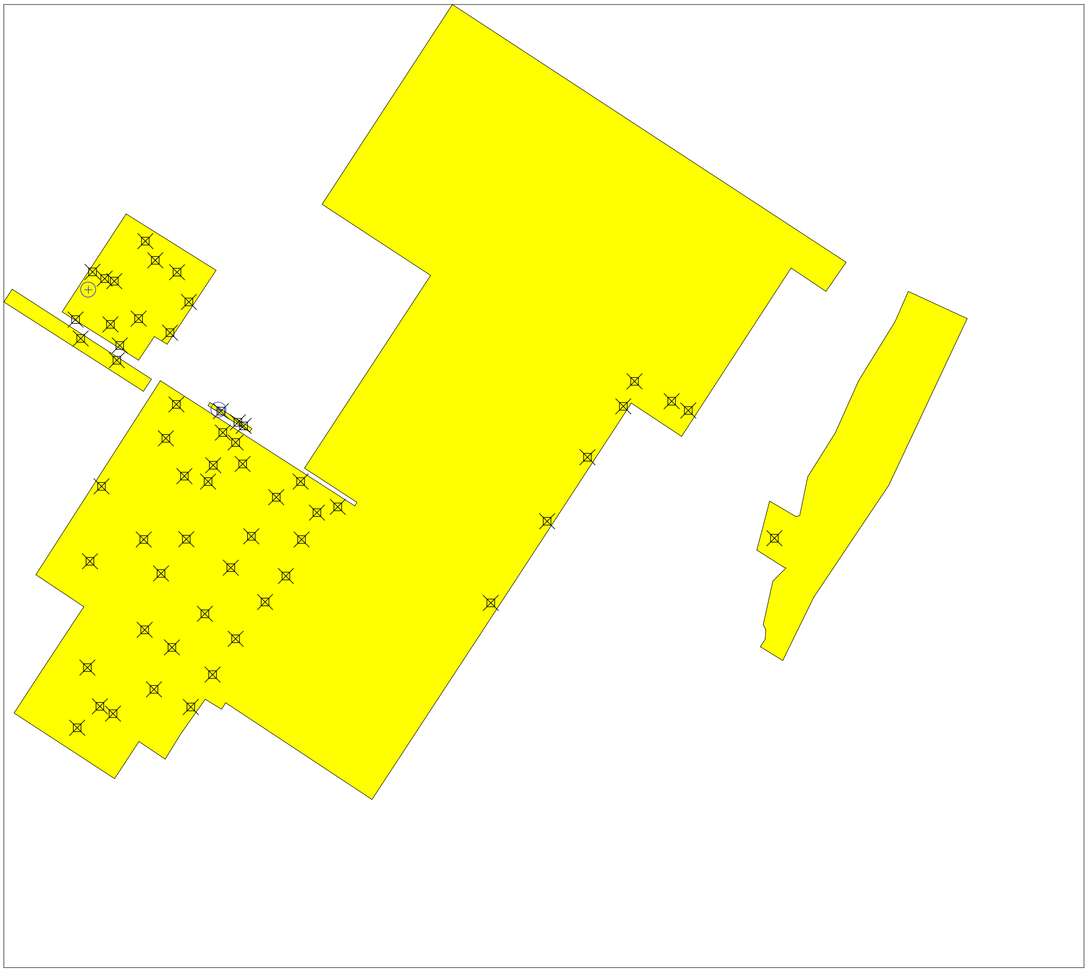
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	Y Coord	Label	Value	Type	Historical
679149.4920	3082933.0980	J-13S	0.54	Manual	T
679279.6830	3083075.4290	J-14S	0.23	Manual	T
679261.0980	3083016.3510	J-15S	0.38	Manual	T
679222.6340	3082840.1720	J-16S	0.33	Manual	T
679293.5600	3082950.4980	J-17S	0.41	Manual	T
679360.5700	3083026.4980	J-18S	0.44	Manual	T
679343.5810	3082969.5980	J-19S	1.7	Manual	T
679382.8640	3083009.1130	J-20S	2.4	Manual	T
679335.0020	3082941.1720	J-21S	0.2	Manual	T
679252.7130	3082781.0290	J-22S	0.74	Manual	T
679297.0010	3082840.6970	J-23S	1.4	Manual	T
679394.8070	3082971.8300	J-24S	0.3	Manual	T
679146.6460	3082549.7640	J-25S	1	Manual	T
679224.5850	3082683.1400	J-26S	0.22	Manual	T
679169.0760	3082537.3510	J-27S	1	Manual	T
679272.0040	3082652.6750	J-28S	1.2	Manual	T

679329.4380	3082711.0960	J-29S	0.58	Manual	T
679374.4420	3082791.3300	J-30S	0.54	Manual	T
679410.1490	3082845.8460	J-31S	0.23	Manual	T
679453.4760	3082914.1150	J-32S	1.1	Manual	T
679495.8840	3082940.9730	J-33S	0.81	Manual	T
679304.6530	3082548.6880	J-34S	0.8	Manual	T
679342.7410	3082605.3190	J-35S	0.78	Manual	T
679382.8900	3082667.5270	J-36S	1.7	Manual	T
679433.9450	3082731.6820	J-37S	2.1	Manual	T
679470.3570	3082776.7350	J-38S	2.2	Manual	T
679497.3310	3082840.3960	J-39S	0.7	Manual	T
679524.3310	3082886.8990	J-40S	0.62	Manual	T
679560.6070	3082897.2580	J-41S	1.7	Manual	T
679924.8150	3082872.3490	J-47S	1.3	Manual	T
679994.9690	3082983.5100	J-48S	1.2	Manual	T
680057.6580	3083072.0750	J-49S	0.69	Manual	T
680077.3540	3083115.5330	J-50S	0.52	Manual	T
679827.1150	3082729.7460	J-51S	0.63	Manual	T
680141.8730	3083080.8800	J-52S	1.1	Manual	T
680170.5600	3083064.6740	J-53S	2.4	Manual	T
679129.3320	3082802.5620	Composite 1	0.22	Manual	T
679124.7500	3082617.3010	Composite 3	1.6	Manual	T
679107.0750	3082512.5600	Composite 4	0.25	Manual	T
679240.6200	3082579.3320	Composite 2	0.44	Manual	T

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.94	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.22	Manual	T
679396.8510	3083038.0640	J-64S	0.24	Manual	T
679386.3850	3083044.5490	J-63S	0.35	Manual	T
679353.3486	3083065.8470	J-65S	0.85	Random	

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.85	Manual	T
679113.1200	3083190.3150	J-66S	0.18	Manual	T

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.28	Manual	T
679104.2450	3083223.2620	J-02S	0.34	Manual	T
679155.0740	3083294.6960	J-03S	0.2	Manual	T
679171.2970	3083289.7960	J-04S	0.23	Manual	T
679225.8560	3083359.9740	J-05S	0.16	Manual	T
679164.8060	3083214.7100	J-06S	0.23	Manual	T
679242.7260	3083326.5280	J-07S	0.47	Manual	T
679181.2750	3083178.2880	J-08S	0.29	Manual	T
679213.7730	3083224.9730	J-09S	0.27	Manual	T
679280.5440	3083305.6810	J-10S	0.27	Manual	T
679268.7700	3083200.3260	J-11S	0.26	Manual	T
679301.1600	3083254.0340	J-12S	0.18	Manual	T
679126.2502	3083275.0474		0	Random	

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5 Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,

Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

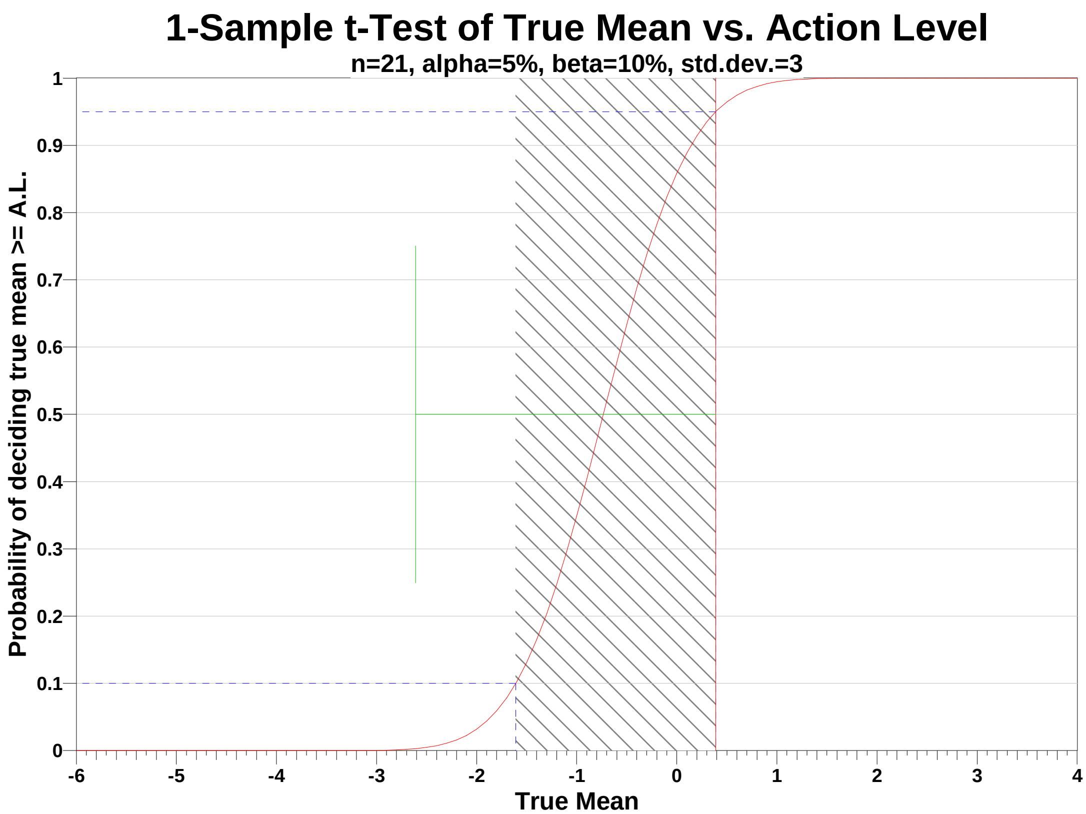
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30

or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.39		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	256148	64038	202696	50675	170162	42541
	β=10	202696	50676	155492	38874	127174	31794
	β=15	170163	42542	127174	31795	101700	25426
LBGR=80	β=5	64038	16011	50675	12670	42541	10636
	β=10	50676	12670	38874	9720	31794	7949
	β=15	42542	10637	31795	7950	25426	6357
LBGR=70	β=5	28463	7117	22523	5632	18908	4728
	β=10	22523	5632	17278	4321	14131	3534
	β=15	18909	4729	14132	3534	11301	2826

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is \$11,500.00, which averages out to a per sample cost of \$547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis
The following data points were entered by the user for analysis.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**

0	0	0.16	0.18	0.18	0.2	0.2	0.22	0.22	0.22	0.23
10	0.23	0.23	0.23	0.24	0.25	0.26	0.27	0.27	0.28	0.29
20	0.3	0.33	0.34	0.35	0.38	0.41	0.44	0.44	0.47	0.52
30	0.54	0.54	0.58	0.62	0.63	0.69	0.7	0.74	0.78	0.8
40	0.81	0.85	0.85	0.94	0.94	1	1	1.1	1.1	1.2
50	1.2	1.3	1.4	1.6	1.7	1.7	1.7	2.1	2.2	2.4
60	2.4									

SUMMARY STATISTICS	
n	61
Min	0
Max	2.4
Range	2.4
Mean	0.72869
Median	0.54
Variance	0.35904
StdDev	0.5992
Std Error	0.076719
Skewness	1.2949
Interquartile Range	0.745

Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.18	0.204	0.255	0.54	1	1.7	2.19	2.4

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	2.781	3.2	No

None of the test statistics exceeded the corresponding critical values, therefore none of the 1 tests are significant and we conclude that at the 5% significance level there are no outliers in the data.

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)

Lilliefors Test Statistic	0.1619
Lilliefors 5% Critical Value	0.1153

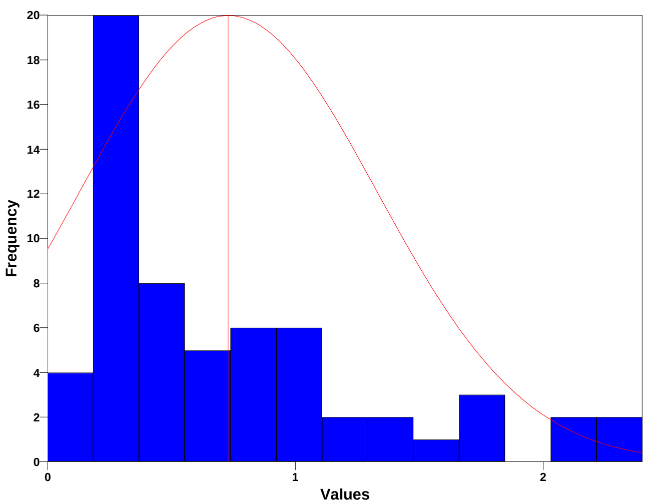
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

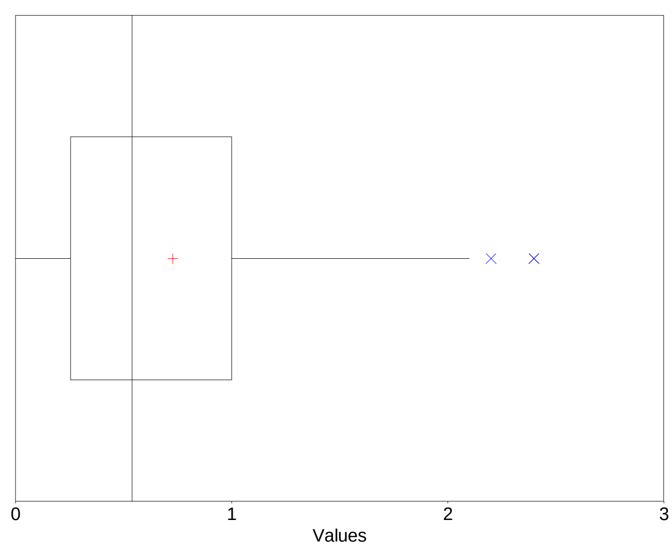
Data Plots
Graphical displays of the data are shown below.

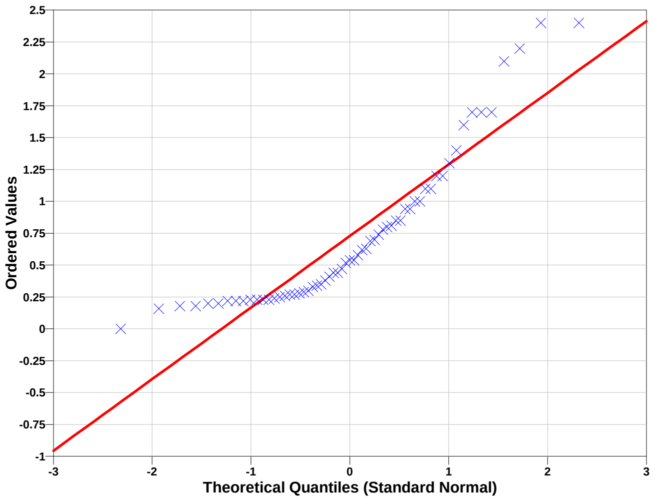
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25th and 75th percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The pth quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.1549
Lilliefors 5% Critical Value	0.1134

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.8569
95% Non-Parametric (Chebyshev) UCL	1.063

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (1.063) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=61 data,
AL is the action level or threshold (0.39),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=60 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
4.4146	1.6706	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
25	37	Cannot Reject
Note: There may not be enough data to reject the null hypothesis (and conclude site is clean) with 95% confidence using the MARSSIM sign test.		

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

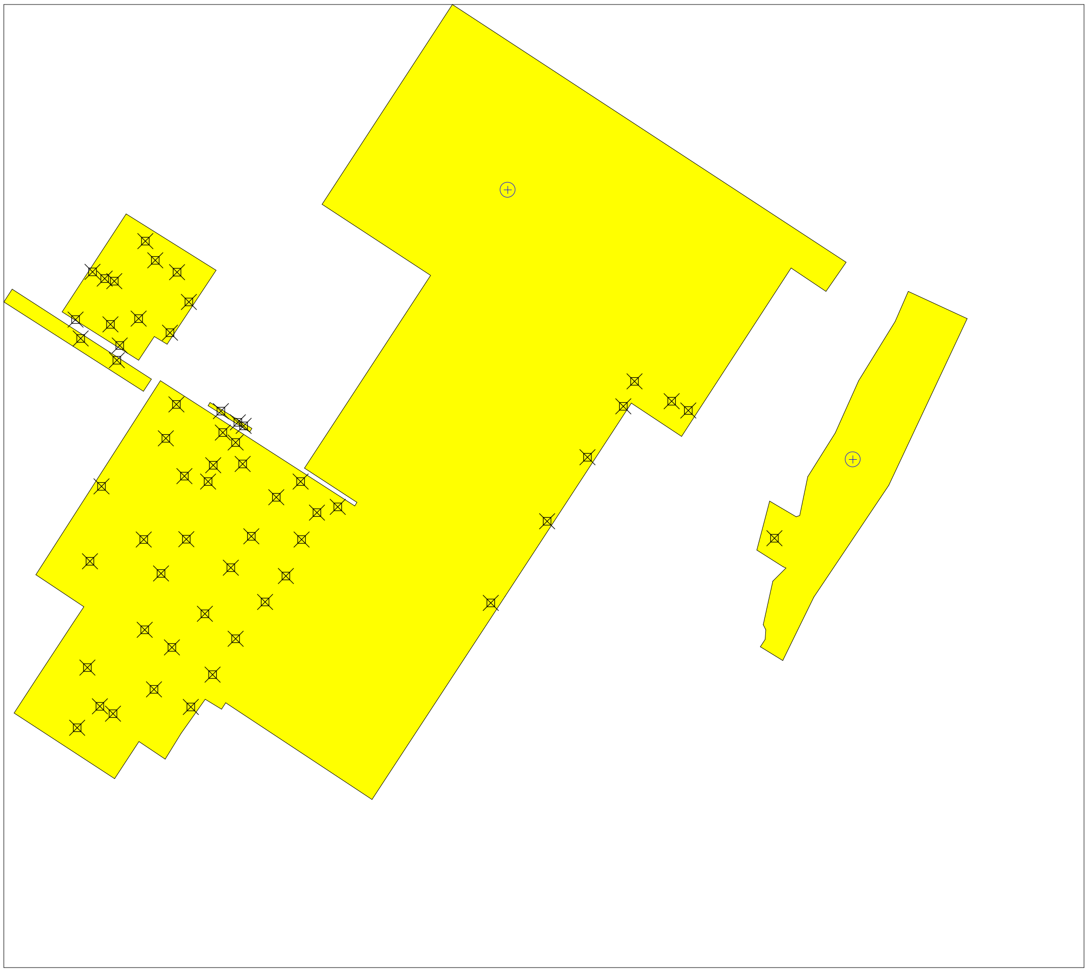
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	Y Coord	Label	Value	Type	Historical
679149.4920	3082933.0980	J-13S	0.54	Manual	T
679279.6830	3083075.4290	J-14S	0.23	Manual	T
679261.0980	3083016.3510	J-15S	0.38	Manual	T
679222.6340	3082840.1720	J-16S	0.33	Manual	T
679293.5600	3082950.4980	J-17S	0.41	Manual	T
679360.5700	3083026.4980	J-18S	0.44	Manual	T
679343.5810	3082969.5980	J-19S	1.7	Manual	T
679382.8640	3083009.1130	J-20S	2.4	Manual	T
679335.0020	3082941.1720	J-21S	0.2	Manual	T
679252.7130	3082781.0290	J-22S	0.74	Manual	T
679297.0010	3082840.6970	J-23S	1.4	Manual	T
679394.8070	3082971.8300	J-24S	0.3	Manual	T
679146.6460	3082549.7640	J-25S	1	Manual	T
679224.5850	3082683.1400	J-26S	0.22	Manual	T
679169.0760	3082537.3510	J-27S	1	Manual	T
679272.0040	3082652.6750	J-28S	1.2	Manual	T

679329.4380	3082711.0960	J-29S	0.58	Manual	T
679374.4420	3082791.3300	J-30S	0.54	Manual	T
679410.1490	3082845.8460	J-31S	0.23	Manual	T
679453.4760	3082914.1150	J-32S	1.1	Manual	T
679495.8840	3082940.9730	J-33S	0.81	Manual	T
679304.6530	3082548.6880	J-34S	0.8	Manual	T
679342.7410	3082605.3190	J-35S	0.78	Manual	T
679382.8900	3082667.5270	J-36S	1.7	Manual	T
679433.9450	3082731.6820	J-37S	2.1	Manual	T
679470.3570	3082776.7350	J-38S	2.2	Manual	T
679497.3310	3082840.3960	J-39S	0.7	Manual	T
679524.3310	3082886.8990	J-40S	0.62	Manual	T
679560.6070	3082897.2580	J-41S	1.7	Manual	T
679924.8150	3082872.3490	J-47S	1.3	Manual	T
679994.9690	3082983.5100	J-48S	1.2	Manual	T
680057.6580	3083072.0750	J-49S	0.69	Manual	T
680077.3540	3083115.5330	J-50S	0.52	Manual	T
679827.1150	3082729.7460	J-51S	0.63	Manual	T
680141.8730	3083080.8800	J-52S	1.1	Manual	T
680170.5600	3083064.6740	J-53S	2.4	Manual	T
679129.3320	3082802.5620	Composite 1	0.22	Manual	T
679124.7500	3082617.3010	Composite 3	1.6	Manual	T
679107.0750	3082512.5600	Composite 4	0.25	Manual	T
679240.6200	3082579.3320	Composite 2	0.44	Manual	T
679856.3696	3083448.7662	Composite 5	0.94	Random	

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.94	Manual	T
680457.3892	3082979.4959	J-62S	0.22	Random	

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.22	Manual	T
679396.8510	3083038.0640	J-64S	0.24	Manual	T
679386.3850	3083044.5490	J-63S	0.35	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.85	Manual	T
679113.1200	3083190.3150	J-66S	0.18	Manual	T

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.28	Manual	T
679104.2450	3083223.2620	J-02S	0.34	Manual	T
679155.0740	3083294.6960	J-03S	0.2	Manual	T
679171.2970	3083289.7960	J-04S	0.23	Manual	T
679225.8560	3083359.9740	J-05S	0.16	Manual	T
679164.8060	3083214.7100	J-06S	0.23	Manual	T
679242.7260	3083326.5280	J-07S	0.47	Manual	T
679181.2750	3083178.2880	J-08S	0.29	Manual	T
679213.7730	3083224.9730	J-09S	0.27	Manual	T
679280.5440	3083305.6810	J-10S	0.27	Manual	T
679268.7700	3083200.3260	J-11S	0.26	Manual	T
679301.1600	3083254.0340	J-12S	0.18	Manual	T

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,

β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

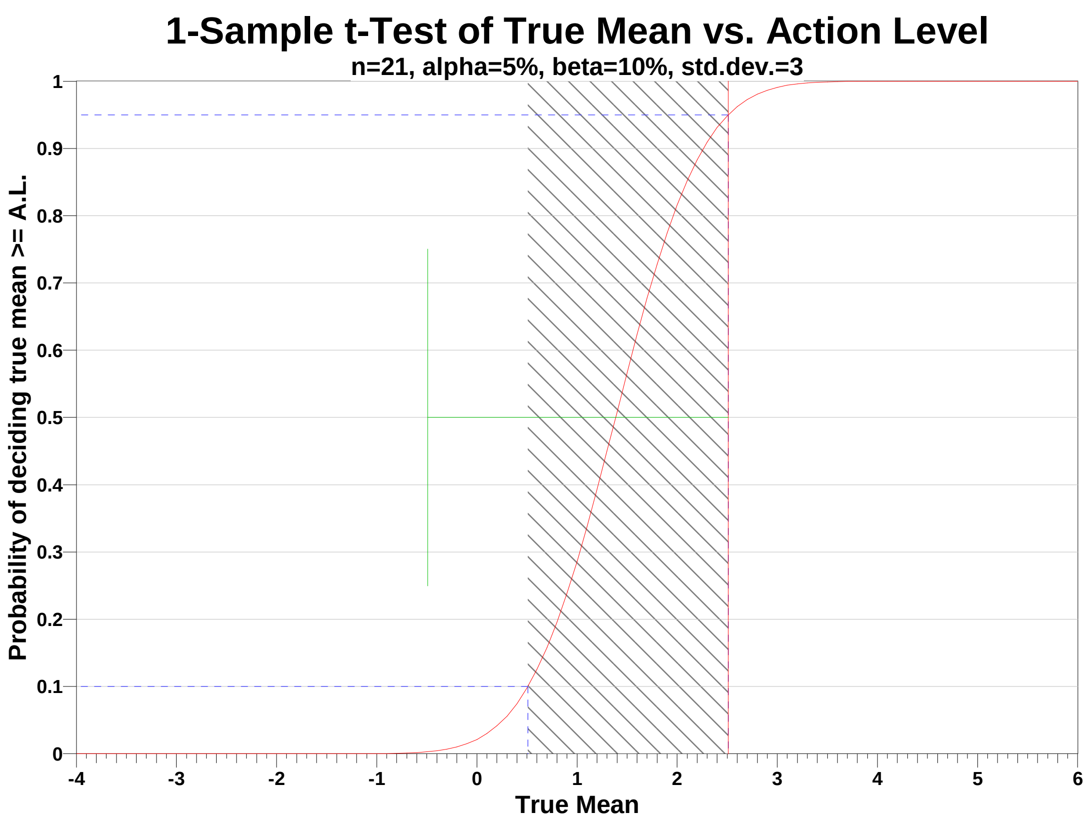
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,

3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=2.50958		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	6188	1548	4897	1225	4111	1028
	β=10	4897	1226	3757	940	3072	769
	β=15	4111	1029	3073	769	2457	615
LBGR=80	β=5	1548	388	1225	307	1028	258
	β=10	1226	308	940	236	769	193
	β=15	1029	259	769	193	615	155
LBGR=70	β=5	689	174	545	137	458	115
	β=10	546	138	419	106	342	86
	β=15	458	116	343	87	274	69

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling

The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10
0	0	0.16	0.18	0.18	0.2	0.2	0.22	0.22	0.22	0.22

10	0.23	0.23	0.23	0.23	0.24	0.25	0.26	0.27	0.27	0.28
20	0.28	0.29	0.3	0.33	0.34	0.35	0.38	0.41	0.44	0.44
30	0.47	0.52	0.54	0.54	0.58	0.62	0.63	0.69	0.7	0.74
40	0.78	0.8	0.81	0.85	0.85	0.94	0.94	1	1	1.1
50	1.1	1.2	1.2	1.3	1.4	1.6	1.7	1.7	1.7	2.1
60	2.2	2.4	2.4							

SUMMARY STATISTICS	
n	63
Min	0
Max	2.4
Range	2.4
Mean	0.71349
Median	0.52
Variance	0.35464
StdDev	0.59552
Std Error	0.075028
Skewness	1.3367
Interquartile Range	0.75

Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.18	0.208	0.25	0.52	1	1.7	2.18	2.4

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	2.832	3.218	No

None of the test statistics exceeded the corresponding critical values, therefore none of the 1 tests are significant and we conclude that at the 5% significance level there are no outliers in the data.

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)

Lilliefors Test Statistic	0.1573
Lilliefors 5% Critical Value	0.1125

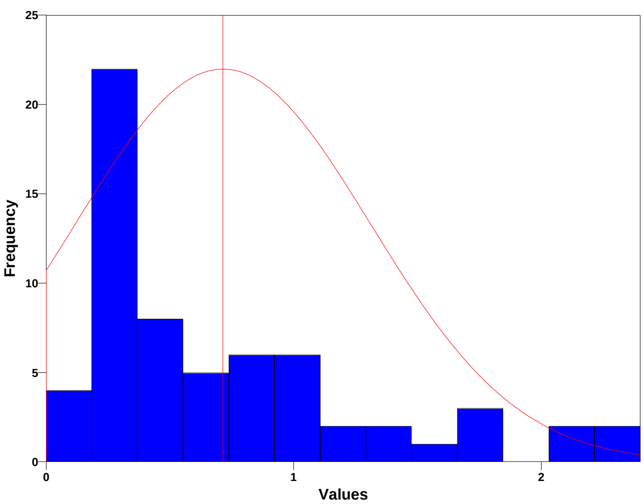
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

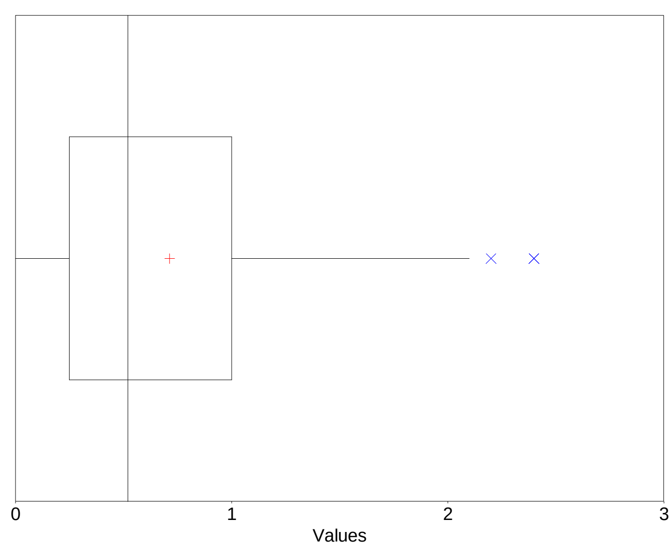
Data Plots
Graphical displays of the data are shown below.

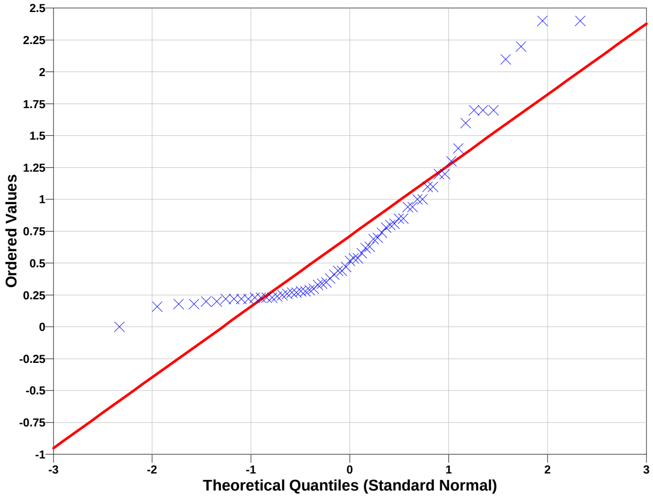
The Histogram is a plot of the fraction of the n observed data that fall within specified data "bins." A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.1605
Lilliefors 5% Critical Value	0.1116

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.8388
95% Non-Parametric (Chebyshev) UCL	1.041

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (1.041) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test

A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where

x is the sample mean of the n=63 data,
AL is the action level or threshold (2.50958),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=62 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-23.939	1.6698	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
63	39	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

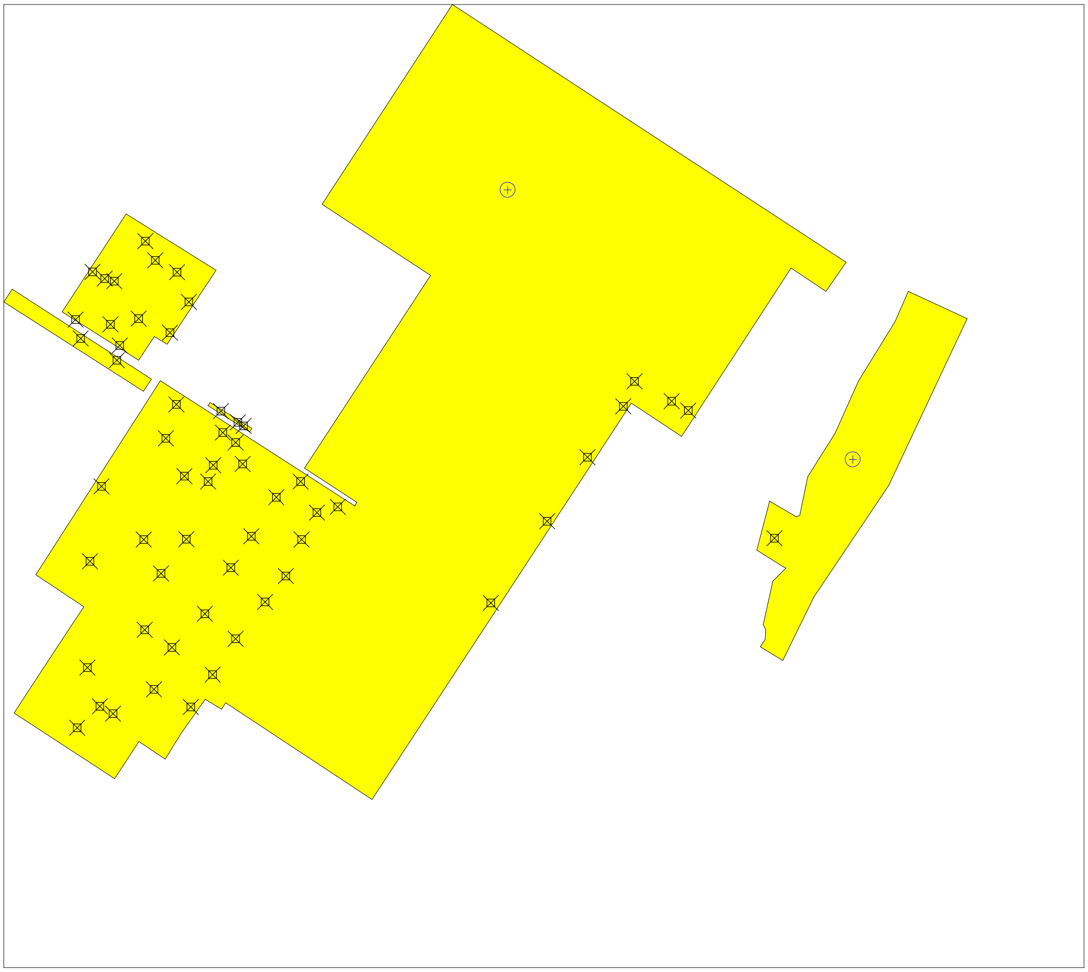
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679149.4920	3082933.0980	J-13S	0.54	Manual	T
679279.6830	3083075.4290	J-14S	0.23	Manual	T
679261.0980	3083016.3510	J-15S	0.38	Manual	T
679222.6340	3082840.1720	J-16S	0.33	Manual	T
679293.5600	3082950.4980	J-17S	0.41	Manual	T
679360.5700	3083026.4980	J-18S	0.44	Manual	T
679343.5810	3082969.5980	J-19S	1.7	Manual	T
679382.8640	3083009.1130	J-20S	2.4	Manual	T
679335.0020	3082941.1720	J-21S	0.2	Manual	T
679252.7130	3082781.0290	J-22S	0.74	Manual	T
679297.0010	3082840.6970	J-23S	1.4	Manual	T
679394.8070	3082971.8300	J-24S	0.3	Manual	T
679146.6460	3082549.7640	J-25S	1	Manual	T
679224.5850	3082683.1400	J-26S	0.22	Manual	T
679169.0760	3082537.3510	J-27S	1	Manual	T
679272.0040	3082652.6750	J-28S	1.2	Manual	T

679329.4380	3082711.0960	J-29S	0.58	Manual	T
679374.4420	3082791.3300	J-30S	0.54	Manual	T
679410.1490	3082845.8460	J-31S	0.23	Manual	T
679453.4760	3082914.1150	J-32S	1.1	Manual	T
679495.8840	3082940.9730	J-33S	0.81	Manual	T
679304.6530	3082548.6880	J-34S	0.8	Manual	T
679342.7410	3082605.3190	J-35S	0.78	Manual	T
679382.8900	3082667.5270	J-36S	1.7	Manual	T
679433.9450	3082731.6820	J-37S	2.1	Manual	T
679470.3570	3082776.7350	J-38S	2.2	Manual	T
679497.3310	3082840.3960	J-39S	0.7	Manual	T
679524.3310	3082886.8990	J-40S	0.62	Manual	T
679560.6070	3082897.2580	J-41S	1.7	Manual	T
679924.8150	3082872.3490	J-47S	1.3	Manual	T
679994.9690	3082983.5100	J-48S	1.2	Manual	T
680057.6580	3083072.0750	J-49S	0.69	Manual	T
680077.3540	3083115.5330	J-50S	0.52	Manual	T
679827.1150	3082729.7460	J-51S	0.63	Manual	T
680141.8730	3083080.8800	J-52S	1.1	Manual	T
680170.5600	3083064.6740	J-53S	2.4	Manual	T
679129.3320	3082802.5620	Composite 1	0.22	Manual	T
679124.7500	3082617.3010	Composite 3	1.6	Manual	T
679107.0750	3082512.5600	Composite 4	0.25	Manual	T
679240.6200	3082579.3320	Composite 2	0.44	Manual	T
679856.3696	3083448.7662	Composite 5	0.94	Random	

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.94	Manual	T
680457.3892	3082979.4959	J-62S	0.22	Random	

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.22	Manual	T
679396.8510	3083038.0640	J-64S	0.24	Manual	T
679386.3850	3083044.5490	J-63S	0.35	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.85	Manual	T
679113.1200	3083190.3150	J-66S	0.18	Manual	T

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.28	Manual	T
679104.2450	3083223.2620	J-02S	0.34	Manual	T
679155.0740	3083294.6960	J-03S	0.2	Manual	T
679171.2970	3083289.7960	J-04S	0.23	Manual	T
679225.8560	3083359.9740	J-05S	0.16	Manual	T
679164.8060	3083214.7100	J-06S	0.23	Manual	T
679242.7260	3083326.5280	J-07S	0.47	Manual	T
679181.2750	3083178.2880	J-08S	0.29	Manual	T
679213.7730	3083224.9730	J-09S	0.27	Manual	T
679280.5440	3083305.6810	J-10S	0.27	Manual	T
679268.7700	3083200.3260	J-11S	0.26	Manual	T
679301.1600	3083254.0340	J-12S	0.18	Manual	T

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,

β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

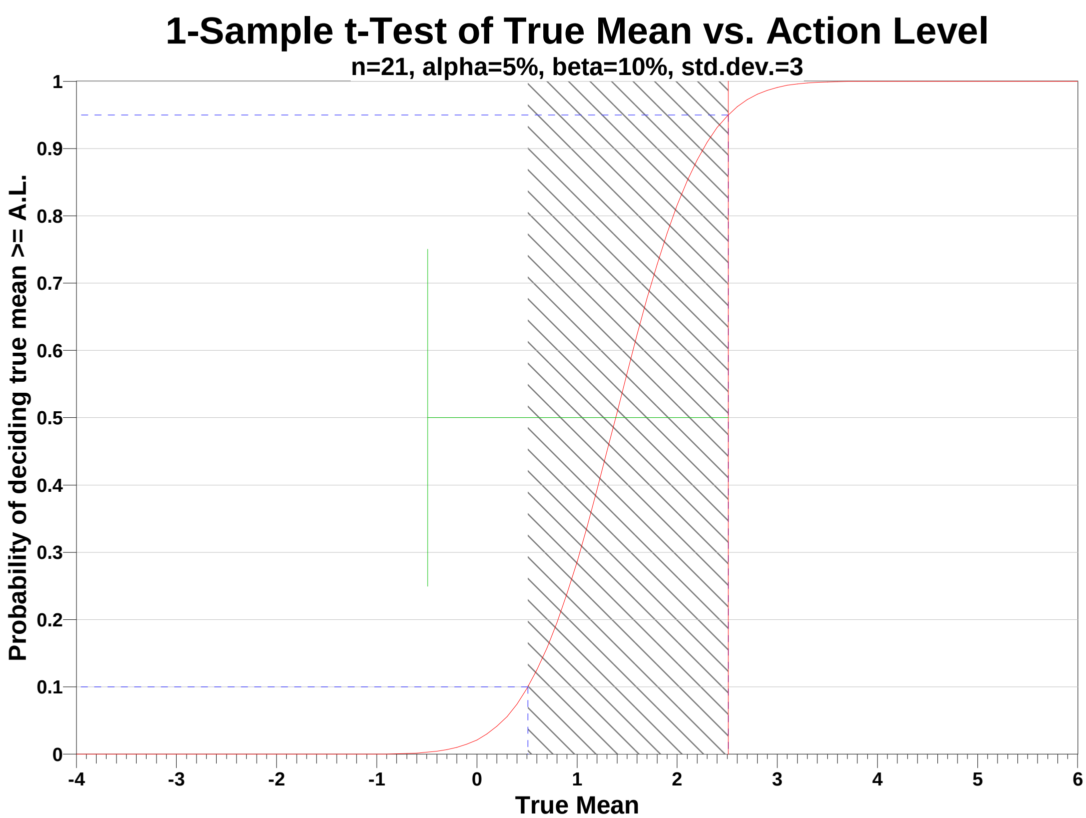
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,

3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=2.50958		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	6188	1548	4897	1225	4111	1028
	β=10	4897	1226	3757	940	3072	769
	β=15	4111	1029	3073	769	2457	615
LBGR=80	β=5	1548	388	1225	307	1028	258
	β=10	1226	308	940	236	769	193
	β=15	1029	259	769	193	615	155
LBGR=70	β=5	689	174	545	137	458	115
	β=10	546	138	419	106	342	86
	β=15	458	116	343	87	274	69

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling

The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10
0	0	0.16	0.18	0.18	0.2	0.2	0.22	0.22	0.22	0.22

10	0.23	0.23	0.23	0.23	0.24	0.25	0.26	0.27	0.27	0.28
20	0.28	0.29	0.3	0.33	0.34	0.35	0.38	0.41	0.44	0.44
30	0.47	0.52	0.54	0.54	0.58	0.62	0.63	0.69	0.7	0.74
40	0.78	0.8	0.81	0.85	0.85	0.94	0.94	1	1	1.1
50	1.1	1.2	1.2	1.3	1.4	1.6	1.7	1.7	1.7	2.1
60	2.2	2.4	2.4							

SUMMARY STATISTICS	
n	63
Min	0
Max	2.4
Range	2.4
Mean	0.71349
Median	0.52
Variance	0.35464
StdDev	0.59552
Std Error	0.075028
Skewness	1.3367
Interquartile Range	0.75

Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.18	0.208	0.25	0.52	1	1.7	2.18	2.4

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	2.832	3.218	No

None of the test statistics exceeded the corresponding critical values, therefore none of the 1 tests are significant and we conclude that at the 5% significance level there are no outliers in the data.

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)

Lilliefors Test Statistic	0.1573
Lilliefors 5% Critical Value	0.1125

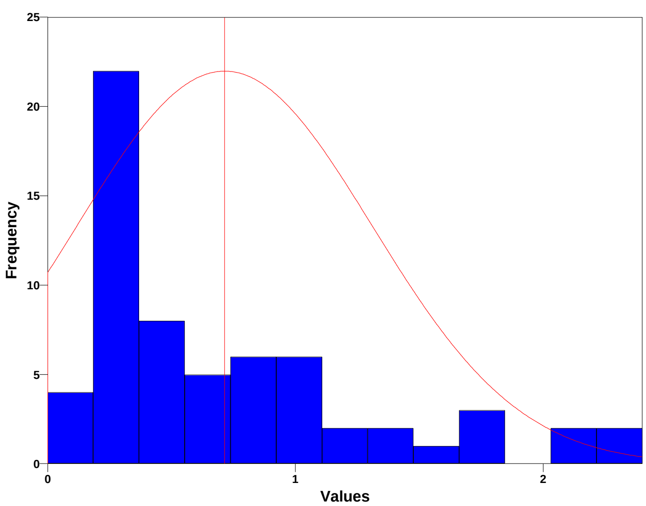
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

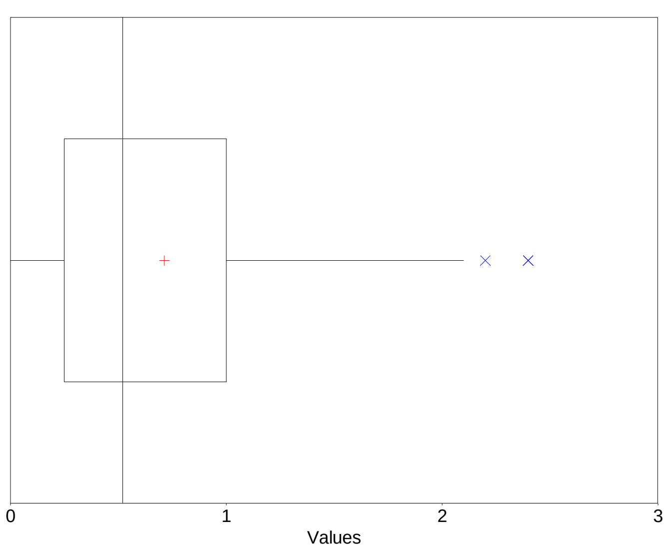
Data Plots
Graphical displays of the data are shown below.

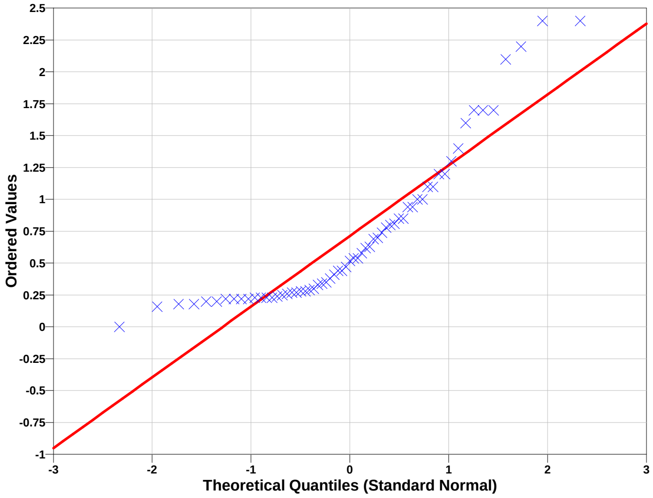
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25th and 75th percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The pth quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.1605
Lilliefors 5% Critical Value	0.1116

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.8388
95% Non-Parametric (Chebyshev) UCL	1.041

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (1.041) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=63 data,
AL is the action level or threshold (2.50958),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=62 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-23.939	1.6698	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
63	39	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

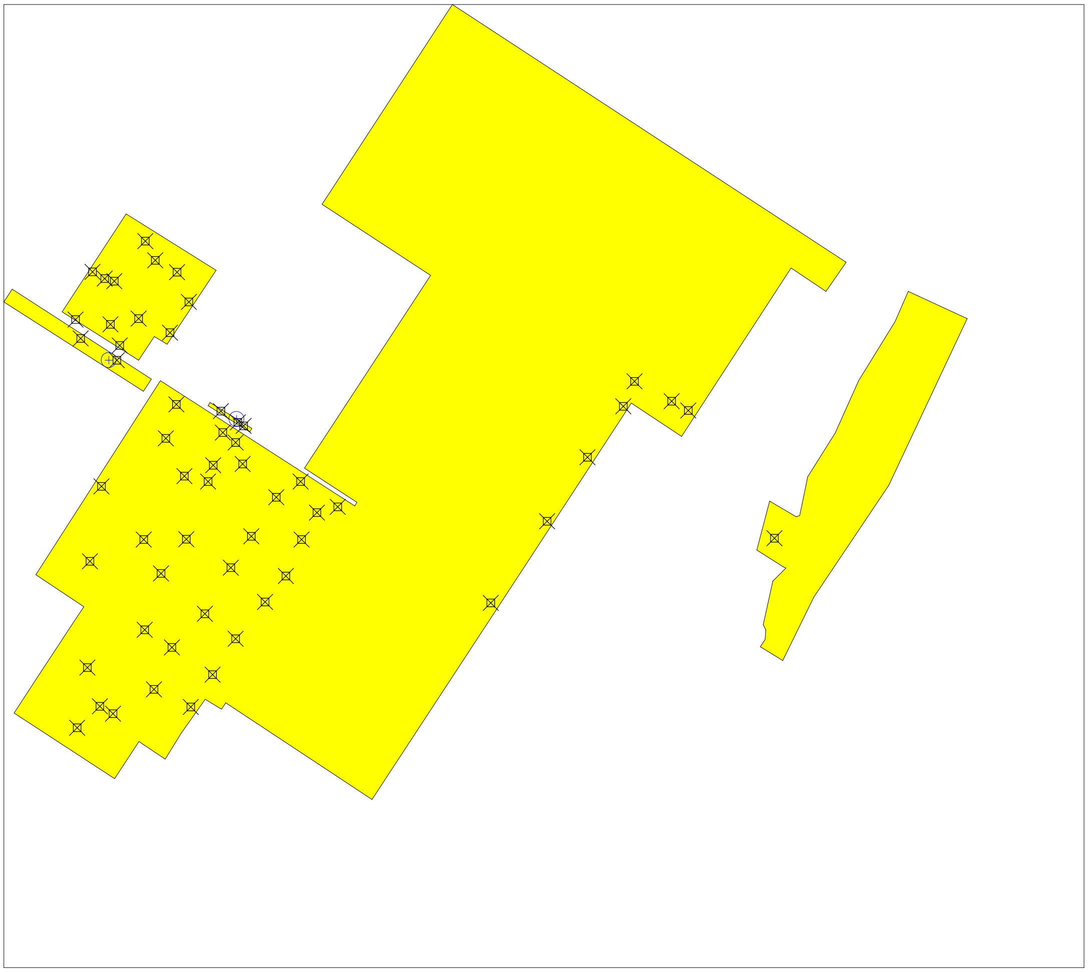
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679149.4920	3082933.0980	J-13S	0.54	Manual	T
679279.6830	3083075.4290	J-14S	0.23	Manual	T
679261.0980	3083016.3510	J-15S	0.38	Manual	T
679222.6340	3082840.1720	J-16S	0.33	Manual	T
679293.5600	3082950.4980	J-17S	0.41	Manual	T
679360.5700	3083026.4980	J-18S	0.44	Manual	T
679343.5810	3082969.5980	J-19S	1.7	Manual	T
679382.8640	3083009.1130	J-20S	2.4	Manual	T
679335.0020	3082941.1720	J-21S	0.2	Manual	T
679252.7130	3082781.0290	J-22S	0.74	Manual	T
679297.0010	3082840.6970	J-23S	1.4	Manual	T
679394.8070	3082971.8300	J-24S	0.3	Manual	T
679146.6460	3082549.7640	J-25S	1	Manual	T
679224.5850	3082683.1400	J-26S	0.22	Manual	T
679169.0760	3082537.3510	J-27S	1	Manual	T
679272.0040	3082652.6750	J-28S	1.2	Manual	T

679329.4380	3082711.0960	J-29S	0.58	Manual	T
679374.4420	3082791.3300	J-30S	0.54	Manual	T
679410.1490	3082845.8460	J-31S	0.23	Manual	T
679453.4760	3082914.1150	J-32S	1.1	Manual	T
679495.8840	3082940.9730	J-33S	0.81	Manual	T
679304.6530	3082548.6880	J-34S	0.8	Manual	T
679342.7410	3082605.3190	J-35S	0.78	Manual	T
679382.8900	3082667.5270	J-36S	1.7	Manual	T
679433.9450	3082731.6820	J-37S	2.1	Manual	T
679470.3570	3082776.7350	J-38S	2.2	Manual	T
679497.3310	3082840.3960	J-39S	0.7	Manual	T
679524.3310	3082886.8990	J-40S	0.62	Manual	T
679560.6070	3082897.2580	J-41S	1.7	Manual	T
679924.8150	3082872.3490	J-47S	1.3	Manual	T
679994.9690	3082983.5100	J-48S	1.2	Manual	T
680057.6580	3083072.0750	J-49S	0.69	Manual	T
680077.3540	3083115.5330	J-50S	0.52	Manual	T
679827.1150	3082729.7460	J-51S	0.63	Manual	T
680141.8730	3083080.8800	J-52S	1.1	Manual	T
680170.5600	3083064.6740	J-53S	2.4	Manual	T
679129.3320	3082802.5620	Composite 1	0.22	Manual	T
679124.7500	3082617.3010	Composite 3	1.6	Manual	T
679107.0750	3082512.5600	Composite 4	0.25	Manual	T
679240.6200	3082579.3320	Composite 2	0.44	Manual	T

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.94	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.22	Manual	T
679396.8510	3083038.0640	J-64S	0.24	Manual	T
679386.3850	3083044.5490	J-63S	0.35	Manual	T
679384.2850	3083049.8773	J-65S	0.85	Random	

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.85	Manual	T
679113.1200	3083190.3150	J-66S	0.18	Manual	T
679161.8021	3083152.5891	J-01S	0.28	Random	

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.28	Manual	T
679104.2450	3083223.2620	J-02S	0.34	Manual	T
679155.0740	3083294.6960	J-03S	0.2	Manual	T
679171.2970	3083289.7960	J-04S	0.23	Manual	T
679225.8560	3083359.9740	J-05S	0.16	Manual	T
679164.8060	3083214.7100	J-06S	0.23	Manual	T
679242.7260	3083326.5280	J-07S	0.47	Manual	T
679181.2750	3083178.2880	J-08S	0.29	Manual	T
679213.7730	3083224.9730	J-09S	0.27	Manual	T
679280.5440	3083305.6810	J-10S	0.27	Manual	T
679268.7700	3083200.3260	J-11S	0.26	Manual	T
679301.1600	3083254.0340	J-12S	0.18	Manual	T

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,

β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.

Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,

3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=24.2		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	68	18	54	14	45	12
	β=10	54	15	42	11	34	9
	β=15	46	13	34	10	27	8
LBGR=80	β=5	18	6	14	5	12	4
	β=10	15	5	11	4	9	3
	β=15	13	5	10	3	8	3
LBGR=70	β=5	9	4	7	3	6	2
	β=10	8	3	6	2	5	2
	β=15	7	3	5	2	4	2

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis
The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10
0	0	0.16	0.18	0.18	0.2	0.2	0.22	0.22	0.22	0.23

10	0.23	0.23	0.23	0.24	0.25	0.26	0.27	0.27	0.28	0.28
20	0.29	0.3	0.33	0.34	0.35	0.38	0.41	0.44	0.44	0.47
30	0.52	0.54	0.54	0.58	0.62	0.63	0.69	0.7	0.74	0.78
40	0.8	0.81	0.85	0.85	0.94	0.94	1	1	1.1	1.1
50	1.2	1.2	1.3	1.4	1.6	1.7	1.7	1.7	2.1	2.2
60	2.4	2.4								

SUMMARY STATISTICS								
n	62							
Min	0							
Max	2.4							
Range	2.4							
Mean	0.72145							
Median	0.53							
Variance	0.3564							
StdDev	0.59699							
Std Error	0.075818							
Skewness	1.3176							
Interquartile Range	0.7425							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.18	0.206	0.2575	0.53	1	1.7	2.185	2.4

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	2.789	3.206	No

None of the test statistics exceeded the corresponding critical values, therefore none of the 1 tests are significant and we conclude that at the 5% significance level there are no outliers in the data.

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)

Lilliefors Test Statistic	0.1517
Lilliefors 5% Critical Value	0.1144

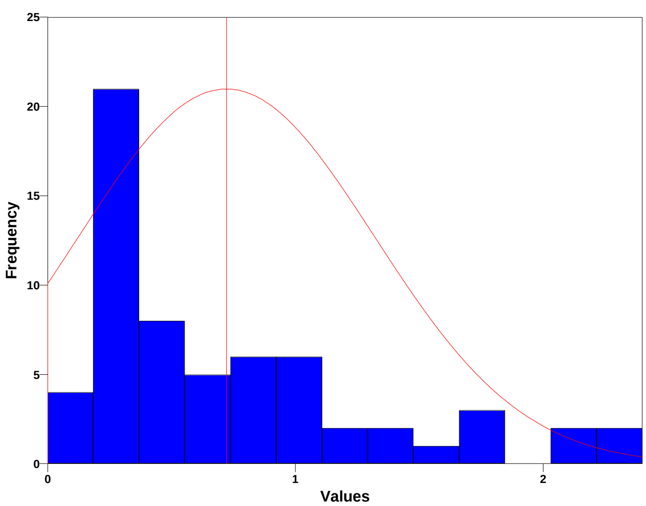
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

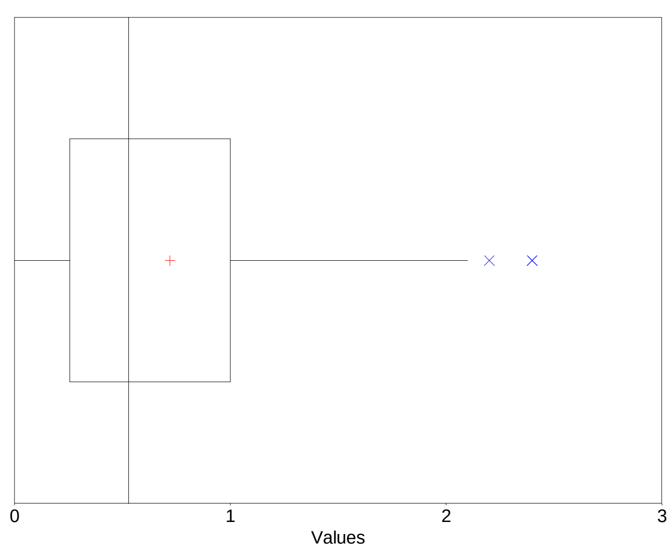
Data Plots
Graphical displays of the data are shown below.

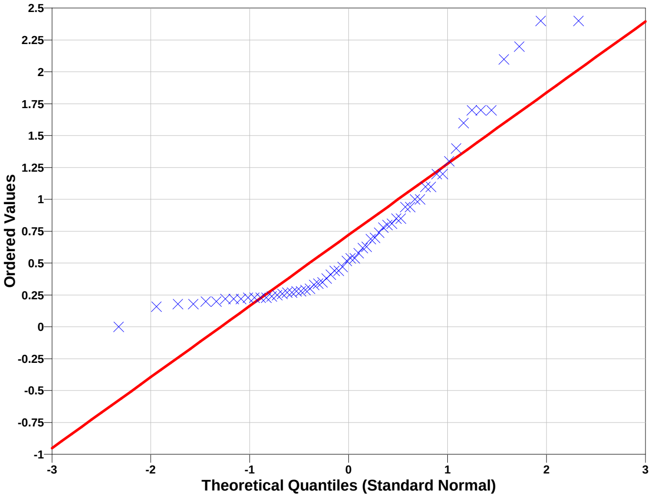
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.1574
Lilliefors 5% Critical Value	0.1125

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.8481
95% Non-Parametric (Chebyshev) UCL	1.052

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (1.052) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test

A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where

x is the sample mean of the n=62 data,
AL is the action level or threshold (24.2),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=61 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-309.67	1.6702	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
62	38	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

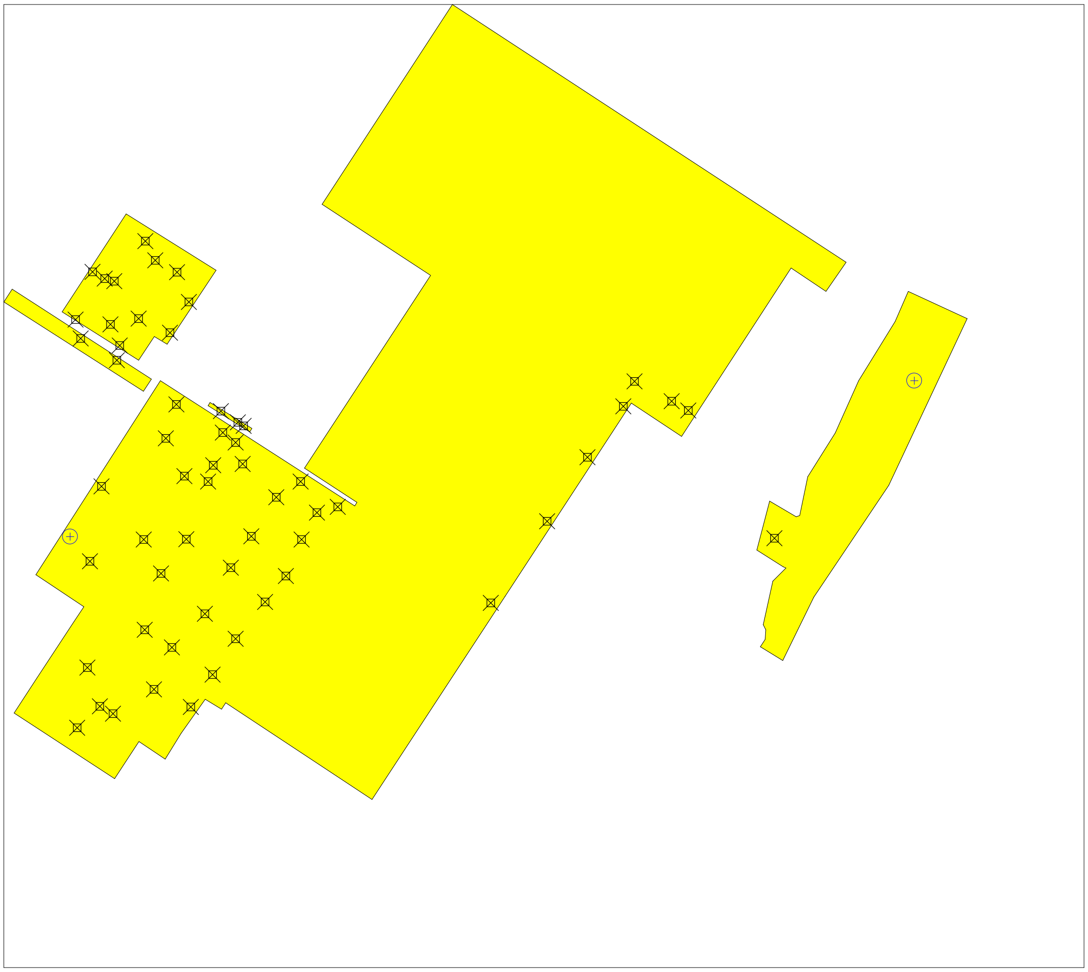
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	Y Coord	Label	Value	Type	Historical
679149.4920	3082933.0980	J-13S	0.54	Manual	T
679279.6830	3083075.4290	J-14S	0.115	Manual	T
679261.0980	3083016.3510	J-15S	0.38	Manual	T
679222.6340	3082840.1720	J-16S	0.33	Manual	T
679293.5600	3082950.4980	J-17S	0.41	Manual	T
679360.5700	3083026.4980	J-18S	0.44	Manual	T
679343.5810	3082969.5980	J-19S	1.7	Manual	T
679382.8640	3083009.1130	J-20S	0.12	Manual	T
679335.0020	3082941.1720	J-21S	0.1	Manual	T
679252.7130	3082781.0290	J-22S	0.74	Manual	T
679297.0010	3082840.6970	J-23S	1.4	Manual	T
679394.8070	3082971.8300	J-24S	0.3	Manual	T
679146.6460	3082549.7640	J-25S	1	Manual	T
679224.5850	3082683.1400	J-26S	0.11	Manual	T
679169.0760	3082537.3510	J-27S	1	Manual	T
679272.0040	3082652.6750	J-28S	1.2	Manual	T

679329.4380	3082711.0960	J-29S	0.58	Manual	T
679374.4420	3082791.3300	J-30S	0.54	Manual	T
679410.1490	3082845.8460	J-31S	0.23	Manual	T
679453.4760	3082914.1150	J-32S	1.1	Manual	T
679495.8840	3082940.9730	J-33S	0.81	Manual	T
679304.6530	3082548.6880	J-34S	0.8	Manual	T
679342.7410	3082605.3190	J-35S	0.78	Manual	T
679382.8900	3082667.5270	J-36S	1.7	Manual	T
679433.9450	3082731.6820	J-37S	2.1	Manual	T
679470.3570	3082776.7350	J-38S	2.2	Manual	T
679497.3310	3082840.3960	J-39S	0.7	Manual	T
679524.3310	3082886.8990	J-40S	0.62	Manual	T
679560.6070	3082897.2580	J-41S	1.7	Manual	T
679924.8150	3082872.3490	J-47S	1.3	Manual	T
679994.9690	3082983.5100	J-48S	1.2	Manual	T
680057.6580	3083072.0750	J-49S	0.69	Manual	T
680077.3540	3083115.5330	J-50S	0.52	Manual	T
679827.1150	3082729.7460	J-51S	0.63	Manual	T
680141.8730	3083080.8800	J-52S	1.1	Manual	T
680170.5600	3083064.6740	J-53S	2.4	Manual	T
679129.3320	3082802.5620	Composite 1	0.11	Manual	T
679240.6200	3082579.3320	Composite 2	1.6	Manual	T
679124.7500	3082617.3010	Composite 3	0.125	Manual	T
679107.0750	3082512.5600	Composite 4	0.44	Manual	T
679094.8663	3082845.4445	Composite 5	0.94	Random	

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.94	Manual	T
680563.9121	3083116.7817	J-62S	0.11	Random	

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.11	Manual	T
679386.3850	3083044.5490	J-63S	0.35	Manual	T
679396.8510	3083038.0640	J-64S	0.12	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.85	Manual	T
679113.1200	3083190.3150	J-66S	0.09	Manual	T

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.14	Manual	T
679104.2450	3083223.2620	J-02S	0.34	Manual	T
679155.0740	3083294.6960	J-03S	0.1	Manual	T
679171.2970	3083289.7960	J-04S	0.115	Manual	T
679225.8560	3083359.9740	J-05S	0.08	Manual	T
679164.8060	3083214.7100	J-06S	0.23	Manual	T
679242.7260	3083326.5280	J-07S	0.47	Manual	T
679181.2750	3083178.2880	J-08S	0.29	Manual	T
679213.7730	3083224.9730	J-09S	0.105	Manual	T
679280.5440	3083305.6810	J-10S	0.105	Manual	T
679268.7700	3083200.3260	J-11S	0.26	Manual	T
679301.1600	3083254.0340	J-12S	0.09	Manual	T

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,

β	is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$	is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is 1-α,
$Z_{1-\beta}$	is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is 1-β.

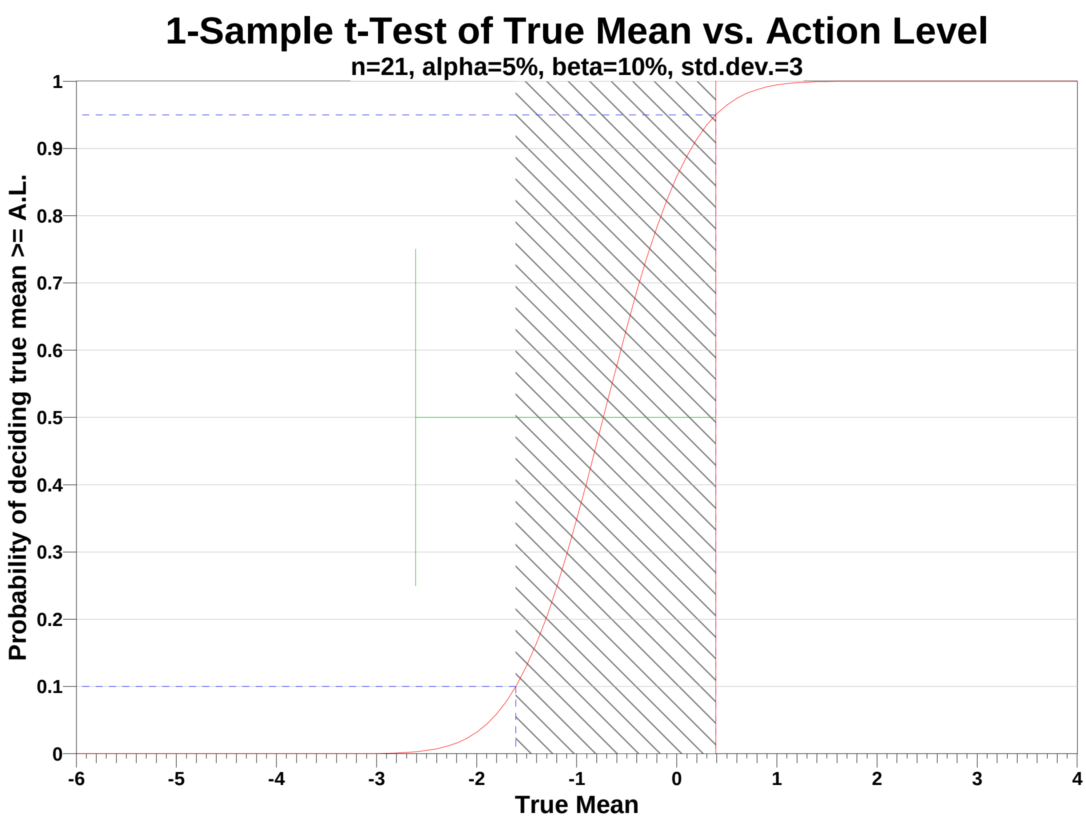
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,

3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.39		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	256148	64038	202696	50675	170162	42541
	β=10	202696	50676	155492	38874	127174	31794
	β=15	170163	42542	127174	31795	101700	25426
LBGR=80	β=5	64038	16011	50675	12670	42541	10636
	β=10	50676	12670	38874	9720	31794	7949
	β=15	42542	10637	31795	7950	25426	6357
LBGR=70	β=5	28463	7117	22523	5632	18908	4728
	β=10	22523	5632	17278	4321	14131	3534
	β=15	18909	4729	14132	3534	11301	2826

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling

The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10
0	0.08	0.09	0.09	0.1	0.1	0.105	0.105	0.11	0.11	0.11

10	0.11	0.115	0.115	0.12	0.12	0.125	0.14	0.23	0.23	0.26
20	0.29	0.3	0.33	0.34	0.35	0.38	0.41	0.44	0.44	0.47
30	0.52	0.54	0.54	0.58	0.62	0.63	0.69	0.7	0.74	0.78
40	0.8	0.81	0.85	0.94	0.94	1	1	1.1	1.1	1.2
50	1.2	1.3	1.4	1.6	1.7	1.7	1.7	2.1	2.2	2.4

SUMMARY STATISTICS	
n	60
Min	0.08
Max	2.4
Range	2.32
Mean	0.66158
Median	0.495
Variance	0.3496
StdDev	0.59127
Std Error	0.076332
Skewness	1.1838
Interquartile Range	0.86375

Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.08	0.0905	0.105	0.1213	0.495	0.985	1.69	2.08	2.4

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	2.909	3.186	No

None of the test statistics exceeded the corresponding critical values, therefore none of the 1 tests are significant and we conclude that at the 5% significance level there are no outliers in the data.

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.1578

Lilliefors 5% Critical Value	0.1174

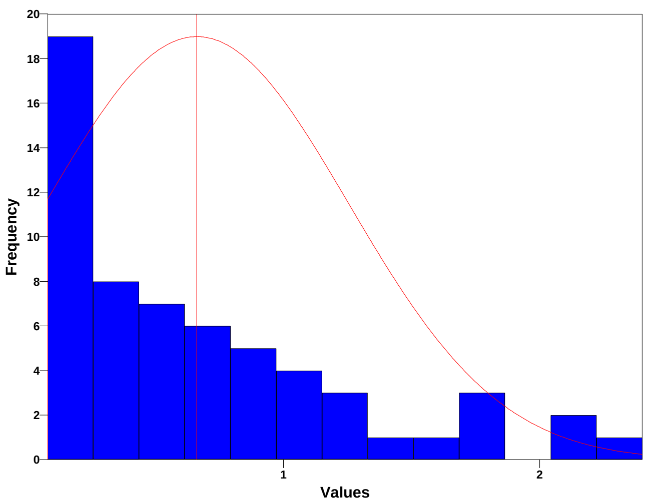
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

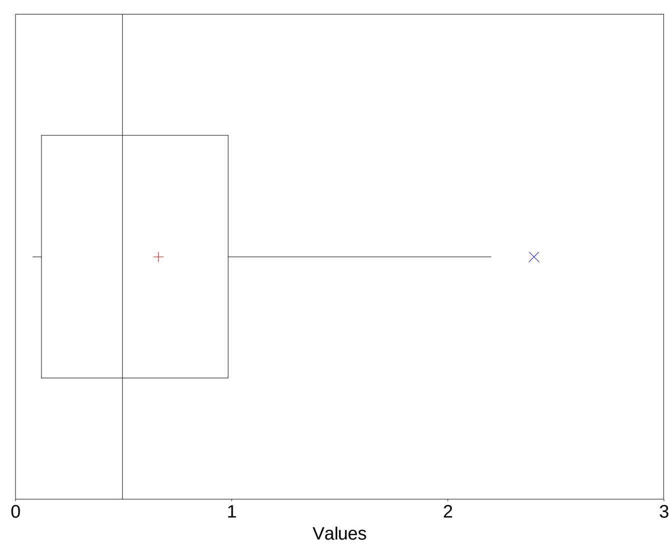
Data Plots
Graphical displays of the data are shown below.

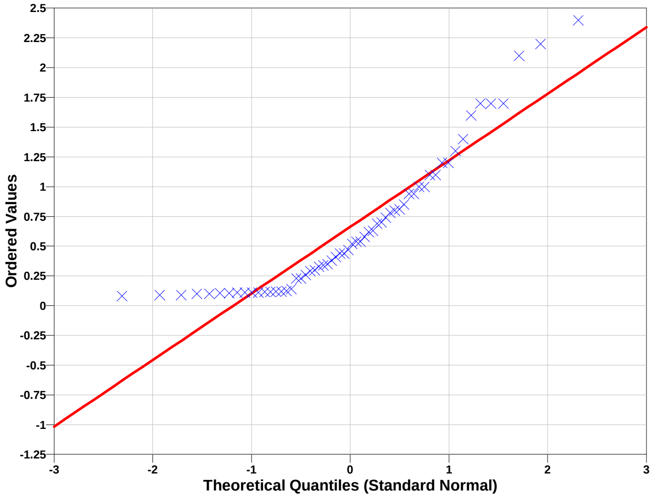
The Histogram is a plot of the fraction of the n observed data that fall within specified data "bins." A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25th and 75th percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.1627
Lilliefors 5% Critical Value	0.1144

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.7891
95% Non-Parametric (Chebyshev) UCL	0.9943

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.9943) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=60 data,
AL is the action level or threshold (0.39),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=59 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
3.5579	1.6711	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
26	37	Cannot Reject
Note: There may not be enough data to reject the null hypothesis (and conclude site is clean) with 95% confidence using the MARSSIM sign test.		

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

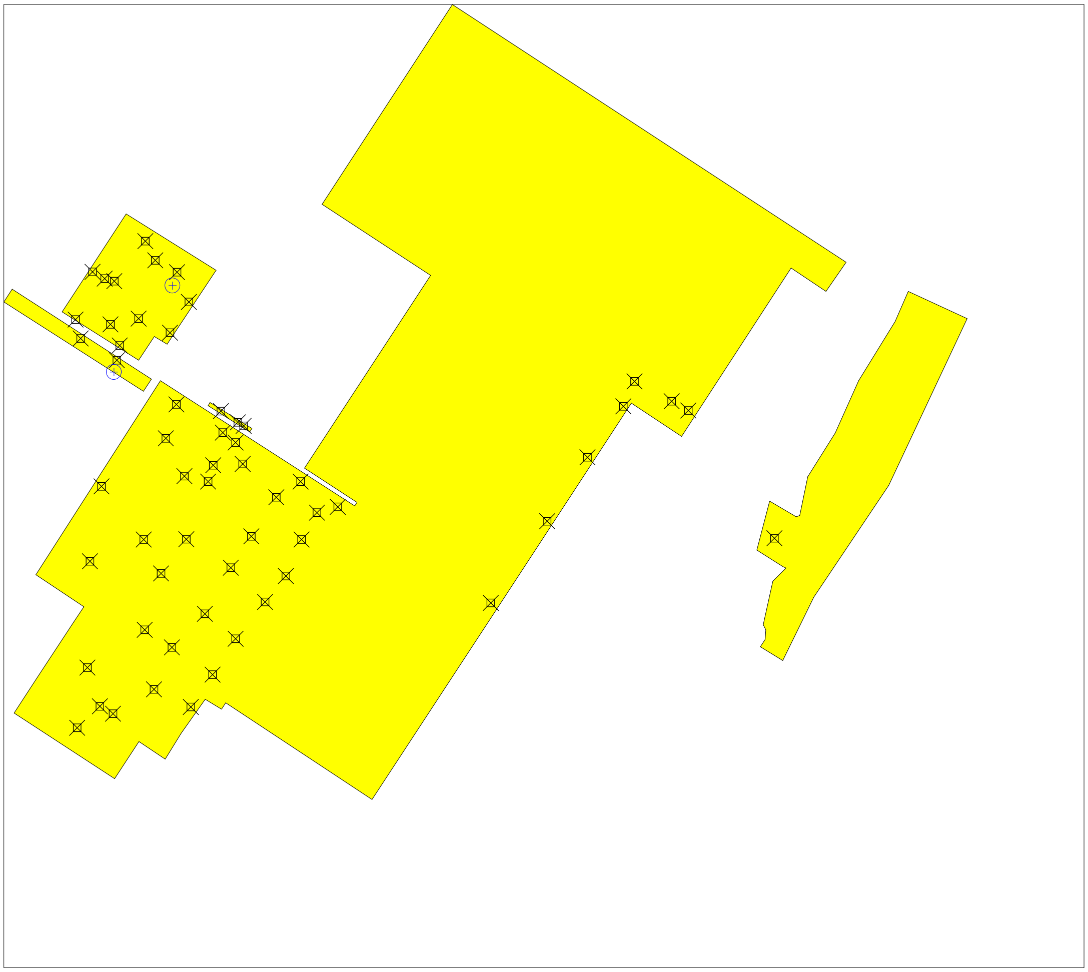
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	Y Coord	Label	Value	Type	Historical
679149.4920	3082933.0980	J-13S	0.077	Manual	T
679279.6830	3083075.4290	J-14S	0.077	Manual	T
679261.0980	3083016.3510	J-15S	0.07	Manual	T
679222.6340	3082840.1720	J-16S	0.072	Manual	T
679293.5600	3082950.4980	J-17S	0.079	Manual	T
679360.5700	3083026.4980	J-18S	0.074	Manual	T
679343.5810	3082969.5980	J-19S	0.077	Manual	T
679382.8640	3083009.1130	J-20S	0.83	Manual	T
679335.0020	3082941.1720	J-21S	0.073	Manual	T
679252.7130	3082781.0290	J-22S	0.072	Manual	T
679297.0010	3082840.6970	J-23S	0.071	Manual	T
679394.8070	3082971.8300	J-24S	0.073	Manual	T
679146.6460	3082549.7640	J-25S	0.064	Manual	T
679224.5850	3082683.1400	J-26S	0.072	Manual	T
679169.0760	3082537.3510	J-27S	0.072	Manual	T
679272.0040	3082652.6750	J-28S	0.073	Manual	T

679329.4380	3082711.0960	J-29S	0.071	Manual	T
679374.4420	3082791.3300	J-30S	0.074	Manual	T
679410.1490	3082845.8460	J-31S	0.079	Manual	T
679453.4760	3082914.1150	J-32S	0.076	Manual	T
679495.8840	3082940.9730	J-33S	0.073	Manual	T
679304.6530	3082548.6880	J-34S	0.073	Manual	T
679342.7410	3082605.3190	J-35S	0.076	Manual	T
679382.8900	3082667.5270	J-36S	0.074	Manual	T
679433.9450	3082731.6820	J-37S	0.071	Manual	T
679470.3570	3082776.7350	J-38S	0.078	Manual	T
679497.3310	3082840.3960	J-39S	0.074	Manual	T
679524.3310	3082886.8990	J-40S	0.074	Manual	T
679560.6070	3082897.2580	J-41S	0.072	Manual	T
679924.8150	3082872.3490	J-47S	0.076	Manual	T
679994.9690	3082983.5100	J-48S	0.076	Manual	T
680057.6580	3083072.0750	J-49S	0.079	Manual	T
680077.3540	3083115.5330	J-50S	0.076	Manual	T
679827.1150	3082729.7460	J-51S	0.074	Manual	T
680141.8730	3083080.8800	J-52S	0.076	Manual	T
680170.5600	3083064.6740	J-53S	0.075	Manual	T
679129.3320	3082802.5620	Composite 1	0.074	Manual	T
679124.7500	3082617.3010	Composite 3	0.076	Manual	T
679107.0750	3082512.5600	Composite 4	0.074	Manual	T
679240.6200	3082579.3320	Composite 2	0.075	Manual	T

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.073	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.073	Manual	T
679396.8510	3083038.0640	J-64S	0.074	Manual	T
679386.3850	3083044.5490	J-63S	0.071	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.066	Manual	T
679113.1200	3083190.3150	J-66S	0.067	Manual	T
679171.2061	3083131.7267	J-01S	0.085	Random	

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.085	Manual	T
679104.2450	3083223.2620	J-02S	0.07	Manual	T
679155.0740	3083294.6960	J-03S	0.082	Manual	T
679171.2970	3083289.7960	J-04S	0.33	Manual	T
679225.8560	3083359.9740	J-05S	0.075	Manual	T
679164.8060	3083214.7100	J-06S	0.07	Manual	T
679242.7260	3083326.5280	J-07S	0.07	Manual	T
679181.2750	3083178.2880	J-08S	0.071	Manual	T
679213.7730	3083224.9730	J-09S	0.072	Manual	T
679280.5440	3083305.6810	J-10S	0.073	Manual	T
679268.7700	3083200.3260	J-11S	0.083	Manual	T
679301.1600	3083254.0340	J-12S	0.072	Manual	T
679273.1230	3083282.5024		0	Random	

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability (1-β) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,

Δ	is the width of the gray region,
α	is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β	is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$	is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is 1-α,
$Z_{1-\beta}$	is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is 1-β.

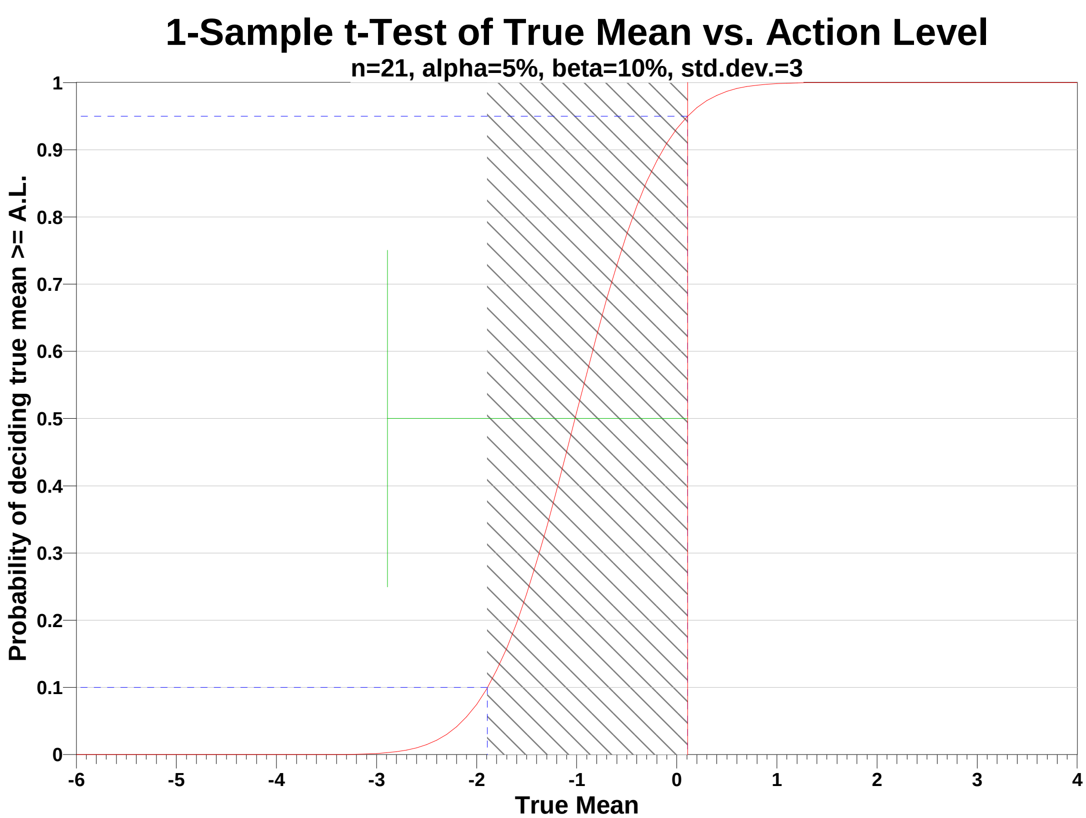
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30

or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.108		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	3340179	835046	2643164	660792	2218921	554731
	β=10	2643165	660793	2027624	506907	1658351	414589
	β=15	2218922	554732	1658351	414589	1326167	331542
LBGR=80	β=5	835046	208763	660792	165199	554731	138684
	β=10	660793	165200	506907	126728	414589	103648
	β=15	554732	138684	414589	103648	331542	82886
LBGR=70	β=5	371133	92785	293686	73423	246548	61638
	β=10	293687	73423	225293	56324	184262	46066
	β=15	246548	61639	184262	46067	147353	36839

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling

The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

The following data points were entered by the user for analysis.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**

0	0	0.064	0.066	0.067	0.07	0.07	0.07	0.07	0.07	0.071
10	0.071	0.071	0.071	0.071	0.072	0.072	0.072	0.072	0.072	0.072
20	0.072	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.074
30	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.075	0.075
40	0.075	0.076	0.076	0.076	0.076	0.076	0.076	0.076	0.077	0.077
50	0.077	0.078	0.079	0.079	0.079	0.082	0.083	0.085	0.085	0.33
60	0.83									

SUMMARY STATISTICS								
n	61							
Min	0							
Max	0.83							
Range	0.83							
Mean	0.08941							
Median	0.074							
Variance	0.010483							
StdDev	0.10239							
Std Error	0.013109							
Skewness	6.7295							
Interquartile Range	0.004							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.0661	0.07	0.072	0.074	0.076	0.0814	0.085	0.83

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	7.205	3.2	Yes

The test statistic 7.205 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	0.83

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.4078
Lilliefors 5% Critical Value	0.1153

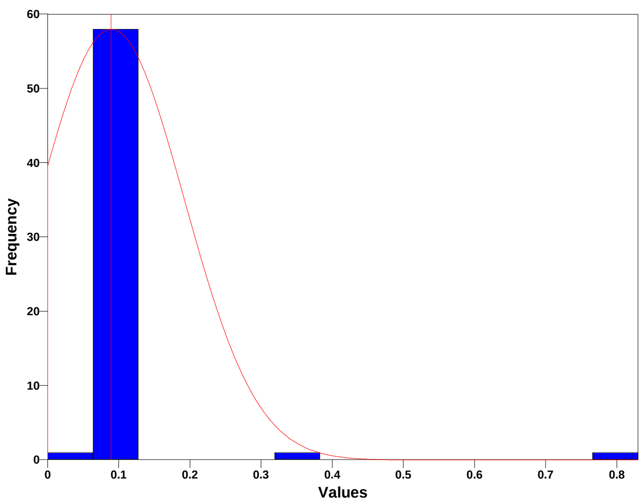
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

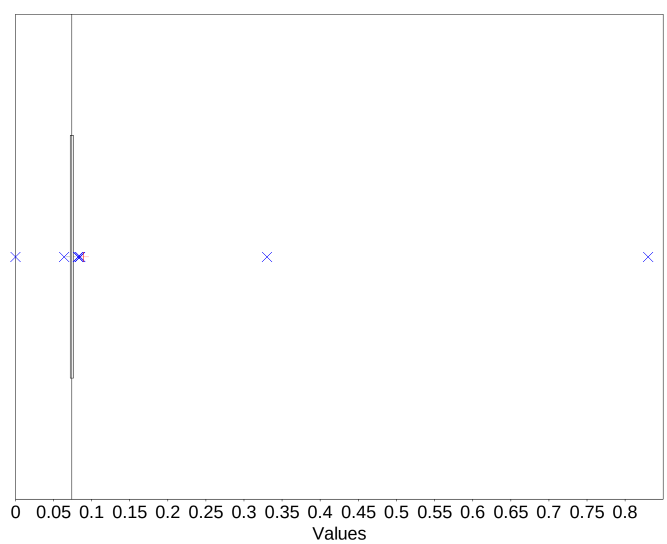
Data Plots
Graphical displays of the data are shown below.

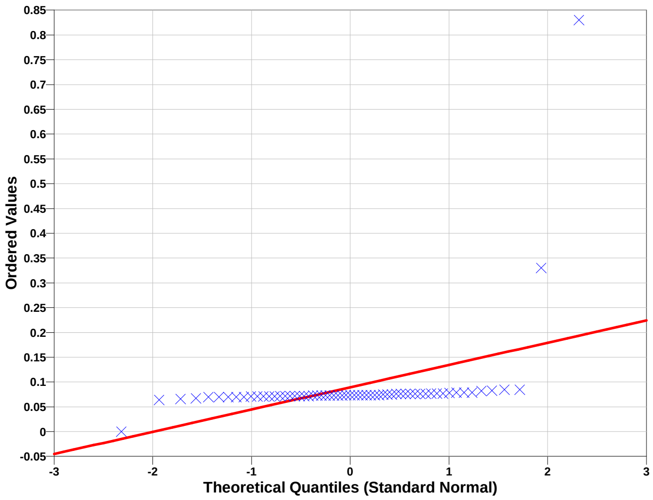
The Histogram is a plot of the fraction of the n observed data that fall within specified data "bins." A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_n, for which a fraction p of the distribution is less than x_n. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.4844
Lilliefors 5% Critical Value	0.1134

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.1113
95% Non-Parametric (Chebyshev) UCL	0.1466

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.1466) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=61 data,
AL is the action level or threshold (0.108),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=60 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-1.4181	1.6706	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
59	37	Reject

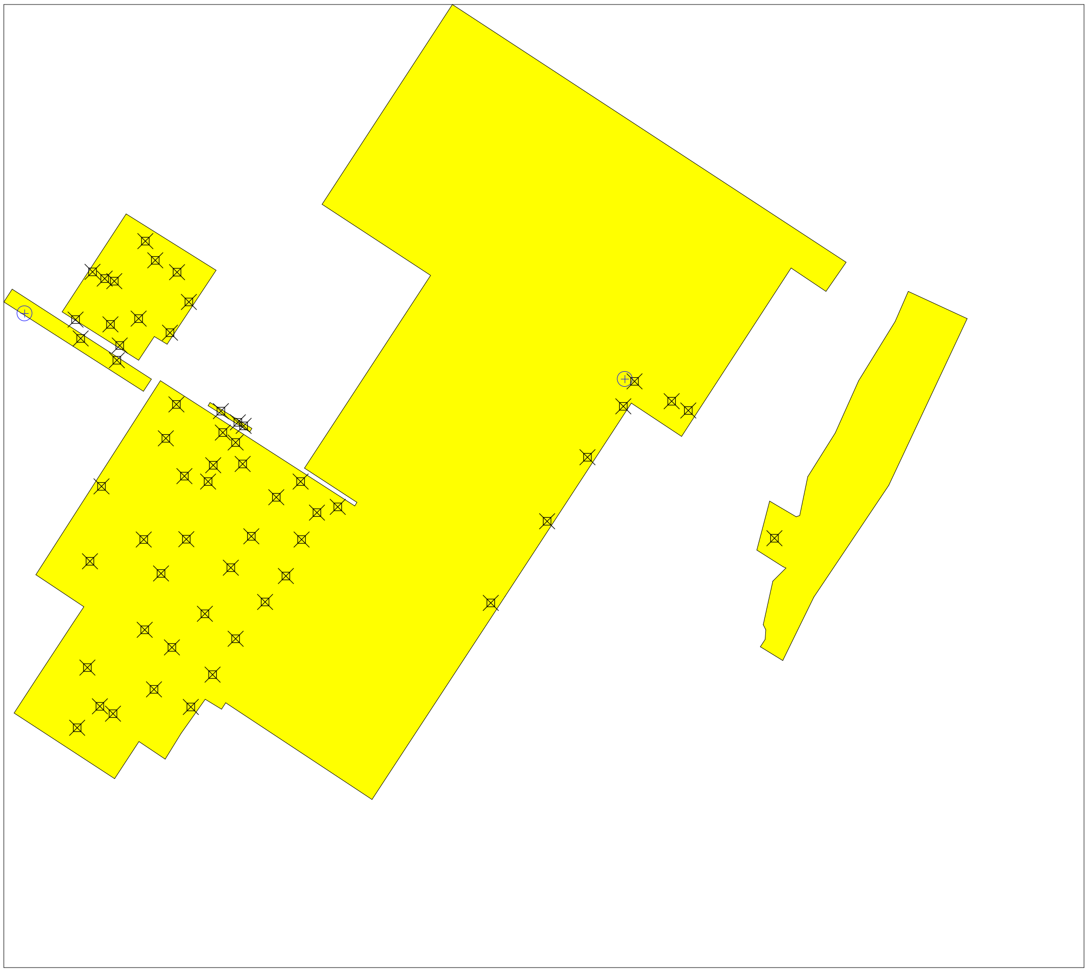
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679149.4920	3082933.0980	J-13S	0.077	Manual	T
679279.6830	3083075.4290	J-14S	0.077	Manual	T
679261.0980	3083016.3510	J-15S	0.07	Manual	T
679222.6340	3082840.1720	J-16S	0.072	Manual	T
679293.5600	3082950.4980	J-17S	0.079	Manual	T
679360.5700	3083026.4980	J-18S	0.074	Manual	T
679343.5810	3082969.5980	J-19S	0.077	Manual	T
679382.8640	3083009.1130	J-20S	0.83	Manual	T
679335.0020	3082941.1720	J-21S	0.073	Manual	T
679252.7130	3082781.0290	J-22S	0.072	Manual	T
679297.0010	3082840.6970	J-23S	0.071	Manual	T
679394.8070	3082971.8300	J-24S	0.073	Manual	T
679146.6460	3082549.7640	J-25S	0.064	Manual	T
679224.5850	3082683.1400	J-26S	0.072	Manual	T
679169.0760	3082537.3510	J-27S	0.072	Manual	T
679272.0040	3082652.6750	J-28S	0.073	Manual	T

679329.4380	3082711.0960	J-29S	0.071	Manual	T
679374.4420	3082791.3300	J-30S	0.074	Manual	T
679410.1490	3082845.8460	J-31S	0.079	Manual	T
679453.4760	3082914.1150	J-32S	0.076	Manual	T
679495.8840	3082940.9730	J-33S	0.073	Manual	T
679304.6530	3082548.6880	J-34S	0.073	Manual	T
679342.7410	3082605.3190	J-35S	0.076	Manual	T
679382.8900	3082667.5270	J-36S	0.074	Manual	T
679433.9450	3082731.6820	J-37S	0.071	Manual	T
679470.3570	3082776.7350	J-38S	0.078	Manual	T
679497.3310	3082840.3960	J-39S	0.074	Manual	T
679524.3310	3082886.8990	J-40S	0.074	Manual	T
679560.6070	3082897.2580	J-41S	0.072	Manual	T
679924.8150	3082872.3490	J-47S	0.076	Manual	T
679994.9690	3082983.5100	J-48S	0.076	Manual	T
680057.6580	3083072.0750	J-49S	0.079	Manual	T
680077.3540	3083115.5330	J-50S	0.076	Manual	T
679827.1150	3082729.7460	J-51S	0.074	Manual	T
680141.8730	3083080.8800	J-52S	0.076	Manual	T
680170.5600	3083064.6740	J-53S	0.075	Manual	T
679129.3320	3082802.5620	Composite 1	0.074	Manual	T
679124.7500	3082617.3010	Composite 3	0.076	Manual	T
679107.0750	3082512.5600	Composite 4	0.074	Manual	T
679240.6200	3082579.3320	Composite 2	0.075	Manual	T
680060.2624	3083119.4017	Composite 5	0.073	Random	

Area: Area 2

X Coord	Y Coord	Label	Value	Type	Historical
680320.6560	3082842.6400	Composite 5	0.073	Manual	T

Area: Area 3

X Coord	Y Coord	Label	Value	Type	Historical
679356.9310	3083064.0350	J-62S	0.073	Manual	T
679396.8510	3083038.0640	J-64S	0.074	Manual	T
679386.3850	3083044.5490	J-63S	0.071	Manual	T

Area: Area 4

X Coord	Y Coord	Label	Value	Type	Historical
679175.7550	3083152.6270	J-65S	0.066	Manual	T
679113.1200	3083190.3150	J-66S	0.067	Manual	T
679015.7816	3083233.7284	J-01S	0.085	Random	

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.085	Manual	T
679104.2450	3083223.2620	J-02S	0.07	Manual	T
679155.0740	3083294.6960	J-03S	0.082	Manual	T
679171.2970	3083289.7960	J-04S	0.33	Manual	T
679225.8560	3083359.9740	J-05S	0.075	Manual	T
679164.8060	3083214.7100	J-06S	0.07	Manual	T
679242.7260	3083326.5280	J-07S	0.07	Manual	T
679181.2750	3083178.2880	J-08S	0.071	Manual	T
679213.7730	3083224.9730	J-09S	0.072	Manual	T
679280.5440	3083305.6810	J-10S	0.073	Manual	T
679268.7700	3083200.3260	J-11S	0.083	Manual	T
679301.1600	3083254.0340	J-12S	0.072	Manual	T

Primary Sampling Objective

The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach

A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs

The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability (1-β) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n is the number of samples,

S is the estimated standard deviation of the measured values including analytical error,

Δ is the width of the gray region,

α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,

β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

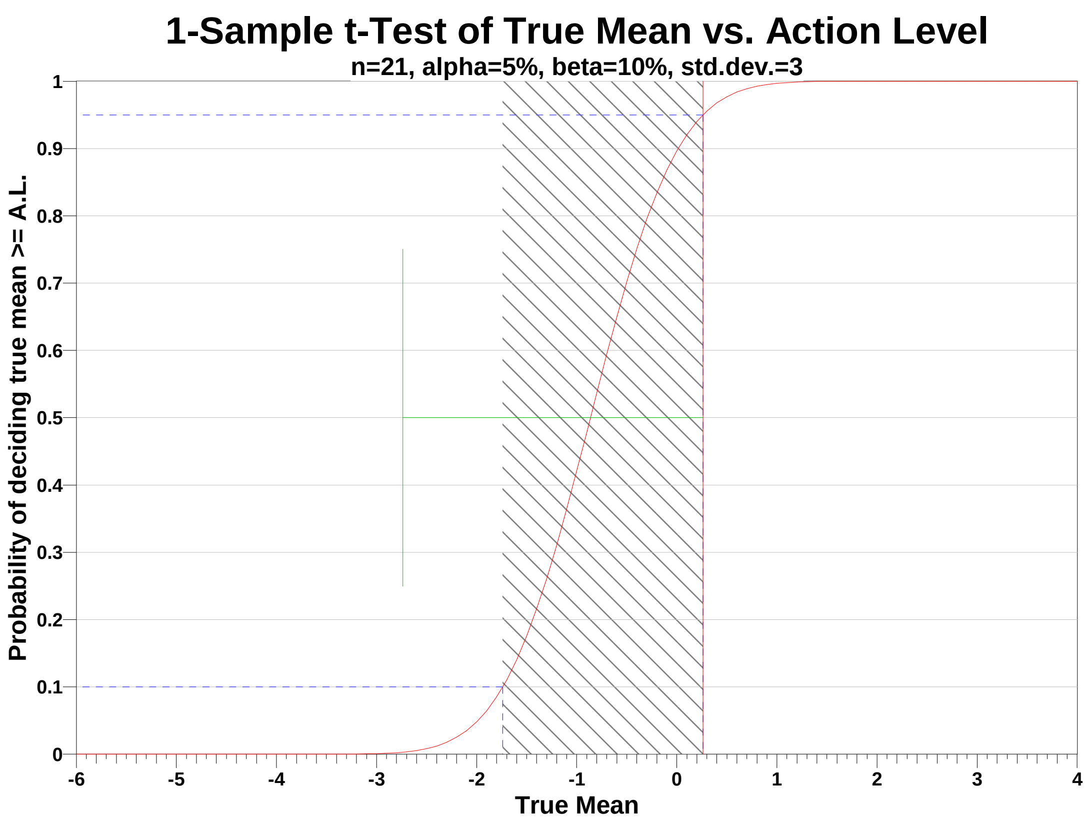
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,

3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.261		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	571923	142982	452576	113145	379935	94985
	β=10	452577	113146	347180	86796	283952	70989
	β=15	379936	94985	283952	70989	227073	56769
LBGR=80	**β=5**	142982	35747	113145	28287	94985	23747
	β=10	113146	28288	86796	21700	70989	17748
	β=15	94985	23748	70989	17748	56769	14193
LBGR=70	**β=5**	63549	15889	50287	12573	42216	10555
	β=10	50288	12573	38577	9645	31551	7889
	β=15	42217	10556	31551	7889	25231	6309

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling

The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

The following data points were entered by the user for analysis.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
0	0	0.064	0.066	0.066	0.067	0.07	0.07	0.07	0.07	0.07

10	0.071	0.071	0.071	0.071	0.071	0.072	0.072	0.072	0.072	0.072
20	0.072	0.072	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073
30	0.073	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.074
40	0.075	0.075	0.075	0.076	0.076	0.076	0.076	0.076	0.076	0.076
50	0.077	0.077	0.077	0.078	0.079	0.079	0.079	0.082	0.083	0.085
60	0.085	0.33	0.83							

SUMMARY STATISTICS	
n	63
Min	0
Max	0.83
Range	0.83
Mean	0.088778
Median	0.074
Variance	0.010157
StdDev	0.10078
Std Error	0.012698
Skewness	6.8396
Interquartile Range	0.004

Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.066	0.07	0.072	0.074	0.076	0.0808	0.085	0.83

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	7.294	3.212	Yes

The test statistic 7.294 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	0.83

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the

suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.3936
Lilliefors 5% Critical Value	0.1134

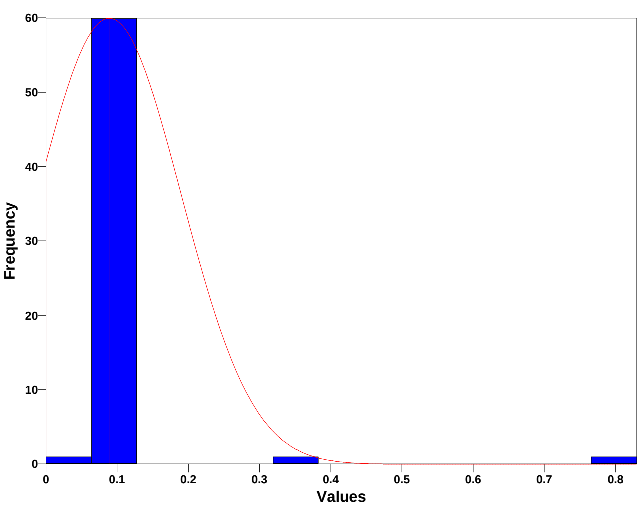
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

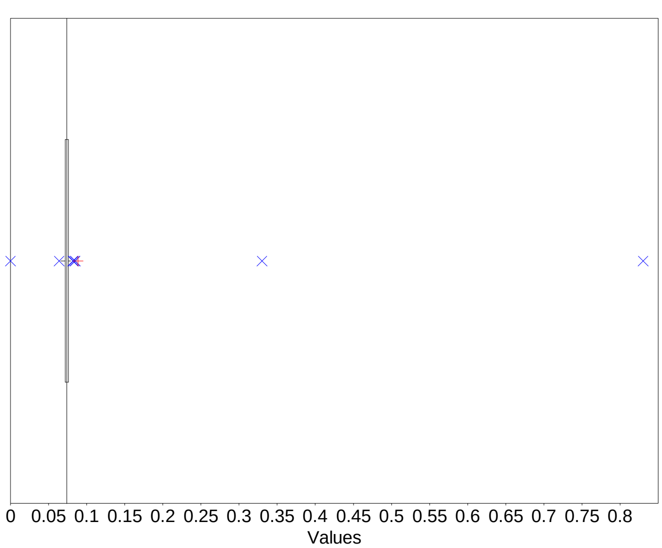
Data Plots
Graphical displays of the data are shown below.

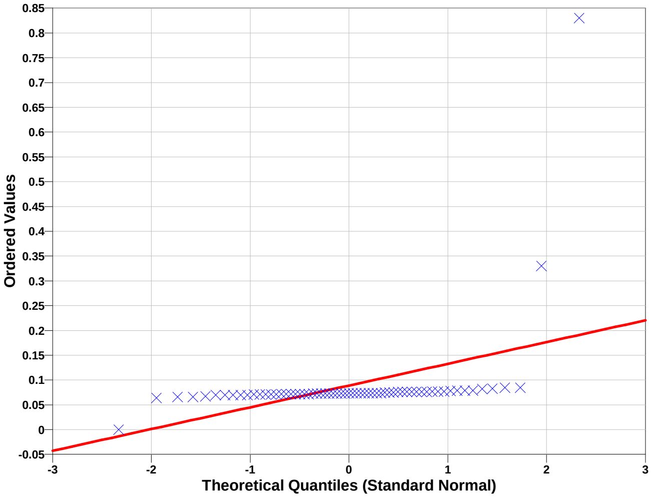
The Histogram is a plot of the fraction of the n observed data that fall within specified data "bins." A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25th and 75th percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The pth quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.4832
Lilliefors 5% Critical Value	0.1116

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.11
95% Non-Parametric (Chebyshev) UCL	0.1441

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.1441) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=63 data,
AL is the action level or threshold (0.261),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=62 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-13.563	1.6698	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
61	39	Reject

Random sampling locations for comparing a mean with a fixed threshold (parametric)

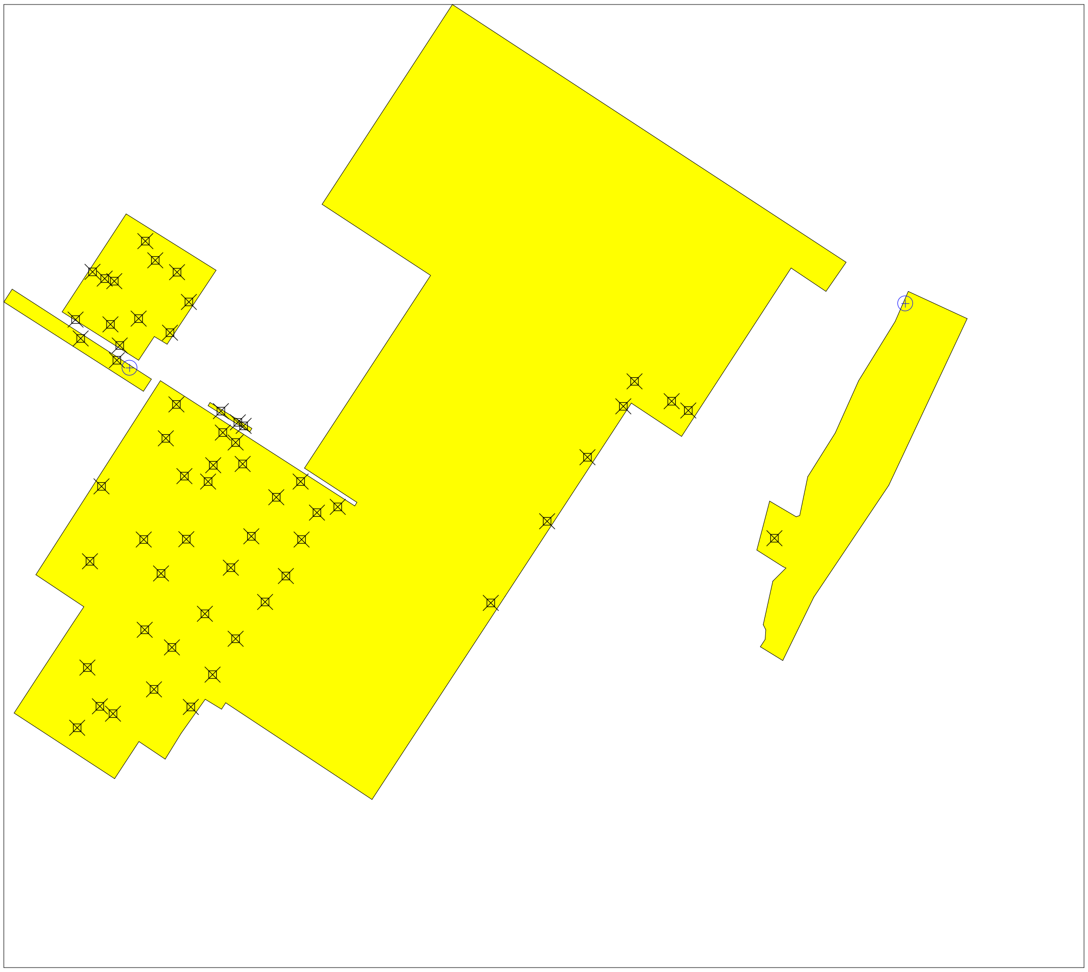
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	Y Coord	Label	Value	Type	Historical
679149.4920	3082933.0980	J-13S	0.077	Manual	T
679279.6830	3083075.4290	J-14S	0.077	Manual	T
679261.0980	3083016.3510	J-15S	0.07	Manual	T
679222.6340	3082840.1720	J-16S	0.072	Manual	T
679293.5600	3082950.4980	J-17S	0.079	Manual	T
679360.5700	3083026.4980	J-18S	0.074	Manual	T
679343.5810	3082969.5980	J-19S	0.077	Manual	T
679382.8640	3083009.1130	J-20S	0.83	Manual	T
679335.0020	3082941.1720	J-21S	0.073	Manual	T
679252.7130	3082781.0290	J-22S	0.072	Manual	T
679297.0010	3082840.6970	J-23S	0.071	Manual	T
679394.8070	3082971.8300	J-24S	0.073	Manual	T
679146.6460	3082549.7640	J-25S	0.064	Manual	T
679224.5850	3082683.1400	J-26S	0.072	Manual	T
679169.0760	3082537.3510	J-27S	0.072	Manual	T
679272.0040	3082652.6750	J-28S	0.073	Manual	T

679329.4380	3082711.0960	J-29S	0.071	Manual	T
679374.4420	3082791.3300	J-30S	0.074	Manual	T
679410.1490	3082845.8460	J-31S	0.079	Manual	T
679453.4760	3082914.1150	J-32S	0.076	Manual	T
679495.8840	3082940.9730	J-33S	0.073	Manual	T
679304.6530	3082548.6880	J-34S	0.073	Manual	T
679342.7410	3082605.3190	J-35S	0.076	Manual	T
679382.8900	3082667.5270	J-36S	0.074	Manual	T
679433.9450	3082731.6820	J-37S	0.071	Manual	T
679470.3570	3082776.7350	J-38S	0.078	Manual	T
679497.3310	3082840.3960	J-39S	0.074	Manual	T
679524.3310	3082886.8990	J-40S	0.074	Manual	T
679560.6070	3082897.2580	J-41S	0.072	Manual	T
679924.8150	3082872.3490	J-47S	0.076	Manual	T
679994.9690	3082983.5100	J-48S	0.076	Manual	T
680057.6580	3083072.0750	J-49S	0.079	Manual	T
680077.3540	3083115.5330	J-50S	0.076	Manual	T
679827.1150	3082729.7460	J-51S	0.074	Manual	T
680141.8730	3083080.8800	J-52S	0.076	Manual	T
680170.5600	3083064.6740	J-53S	0.075	Manual	T
679129.3320	3082802.5620	Composite 1	0.074	Manual	T
679124.7500	3082617.3010	Composite 3	0.076	Manual	T
679107.0750	3082512.5600	Composite 4	0.074	Manual	T
679240.6200	3082579.3320	Composite 2	0.075	Manual	T

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.073	Manual	T
680548.6718	3083251.2433	J-62S	0.073	Random	

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.073	Manual	T
679396.8510	3083038.0640	J-64S	0.074	Manual	T
679386.3850	3083044.5490	J-63S	0.071	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.066	Manual	T
679113.1200	3083190.3150	J-66S	0.067	Manual	T
679198.3689	3083138.7797	J-01S	0.085	Random	

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.085	Manual	T
679104.2450	3083223.2620	J-02S	0.07	Manual	T
679155.0740	3083294.6960	J-03S	0.082	Manual	T
679171.2970	3083289.7960	J-04S	0.33	Manual	T
679225.8560	3083359.9740	J-05S	0.075	Manual	T
679164.8060	3083214.7100	J-06S	0.07	Manual	T
679242.7260	3083326.5280	J-07S	0.07	Manual	T
679181.2750	3083178.2880	J-08S	0.071	Manual	T
679213.7730	3083224.9730	J-09S	0.072	Manual	T
679280.5440	3083305.6810	J-10S	0.073	Manual	T
679268.7700	3083200.3260	J-11S	0.083	Manual	T
679301.1600	3083254.0340	J-12S	0.072	Manual	T

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability (1-β) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5 Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,

β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

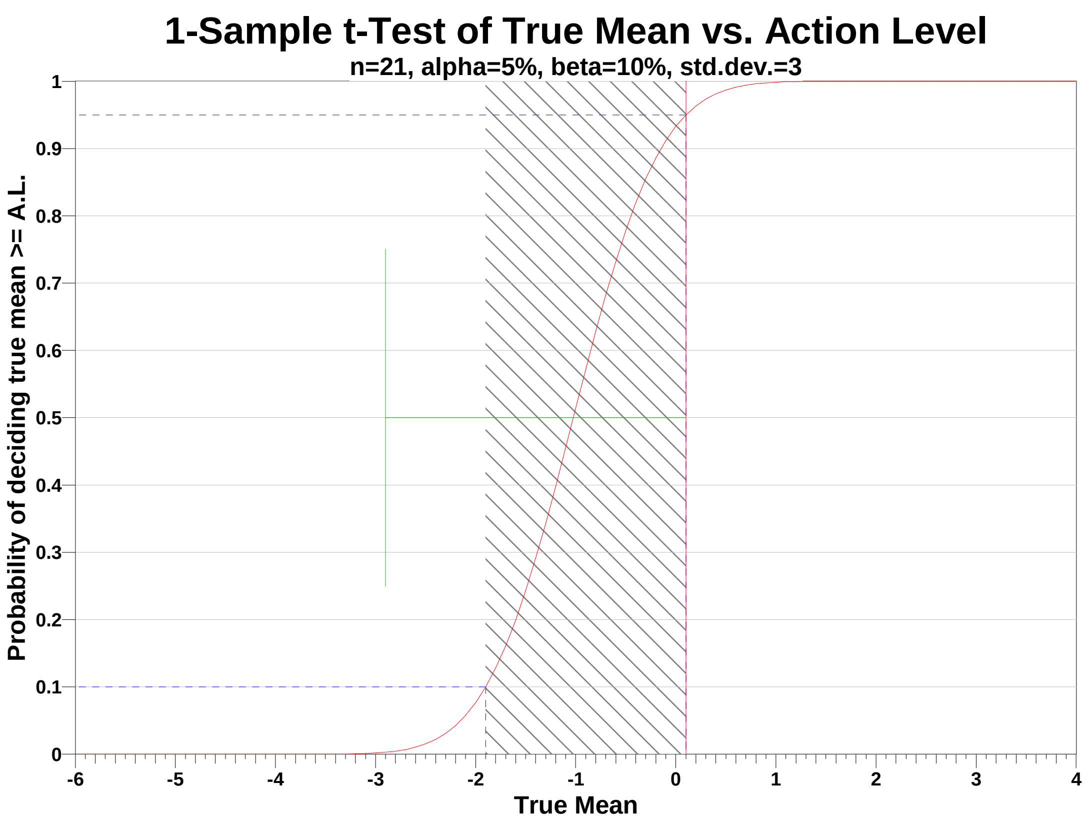
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,

3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.1		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	3895984	973997	3082986	770748	2588149	647038
	β=10	3082987	770748	2365020	591256	1934301	483576
	β=15	2588150	647039	1934301	483576	1546841	386711
LBGR=80	**β=5**	973997	243501	770748	192688	647038	161760
	β=10	770748	192688	591256	147815	483576	120895
	β=15	647039	161761	483576	120895	386711	96679
LBGR=70	**β=5**	432889	108224	342555	85640	287573	71894
	β=10	342556	85640	262781	65696	214923	53732
	β=15	287574	71895	214923	53732	171872	42969

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling

The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

The following data points were entered by the user for analysis.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
0	0.064	0.066	0.067	0.07	0.07	0.07	0.07	0.07	0.071	0.071

10	0.071	0.071	0.071	0.072	0.072	0.072	0.072	0.072	0.072	0.072
20	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.074
30	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.075	0.075
40	0.075	0.076	0.076	0.076	0.076	0.076	0.076	0.076	0.077	0.077
50	0.077	0.078	0.079	0.079	0.079	0.082	0.083	0.085	0.085	0.33
60	0.83									

SUMMARY STATISTICS	
n	61
Min	0.064
Max	0.83
Range	0.766
Mean	0.090607
Median	0.074
Variance	0.010353
StdDev	0.10175
Std Error	0.013028
Skewness	6.832
Interquartile Range	0.004

Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.064	0.0673	0.07	0.072	0.074	0.076	0.0814	0.085	0.83

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	7.205	3.2	Yes

The test statistic 7.205 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	0.83

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the

suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.4078
Lilliefors 5% Critical Value	0.1153

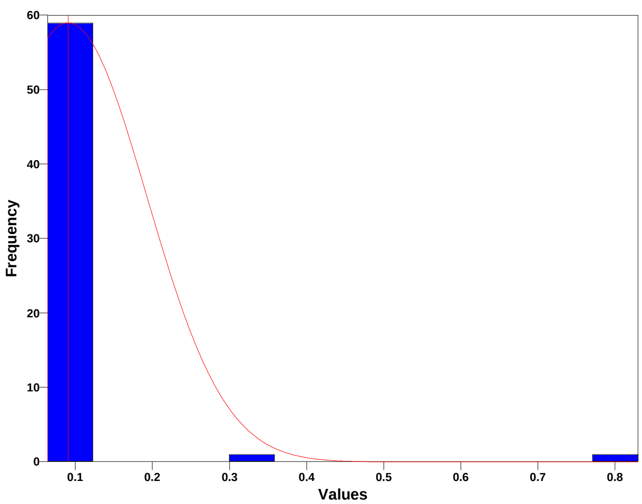
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

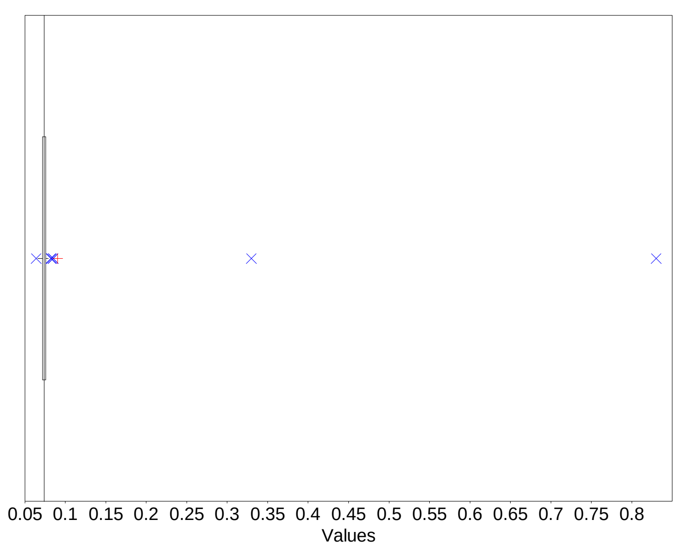
Data Plots
Graphical displays of the data are shown below.

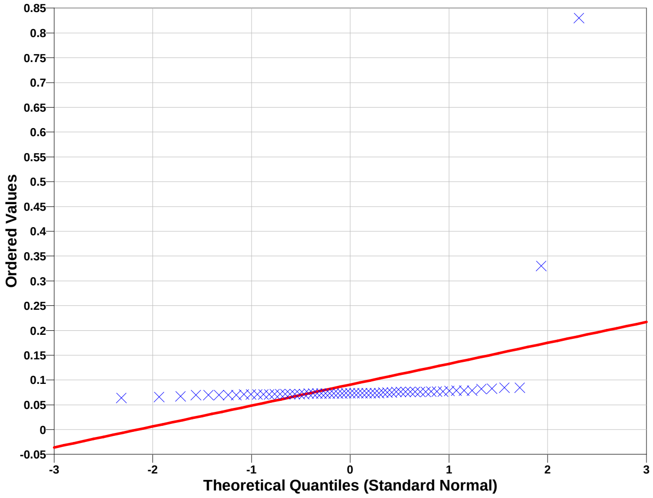
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25th and 75th percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The pth quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.4892
Lilliefors 5% Critical Value	0.1134

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.1124
95% Non-Parametric (Chebyshev) UCL	0.1474

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.1474) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=61 data,
AL is the action level or threshold (0.1),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=60 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-0.72105	1.6706	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
59	37	Reject

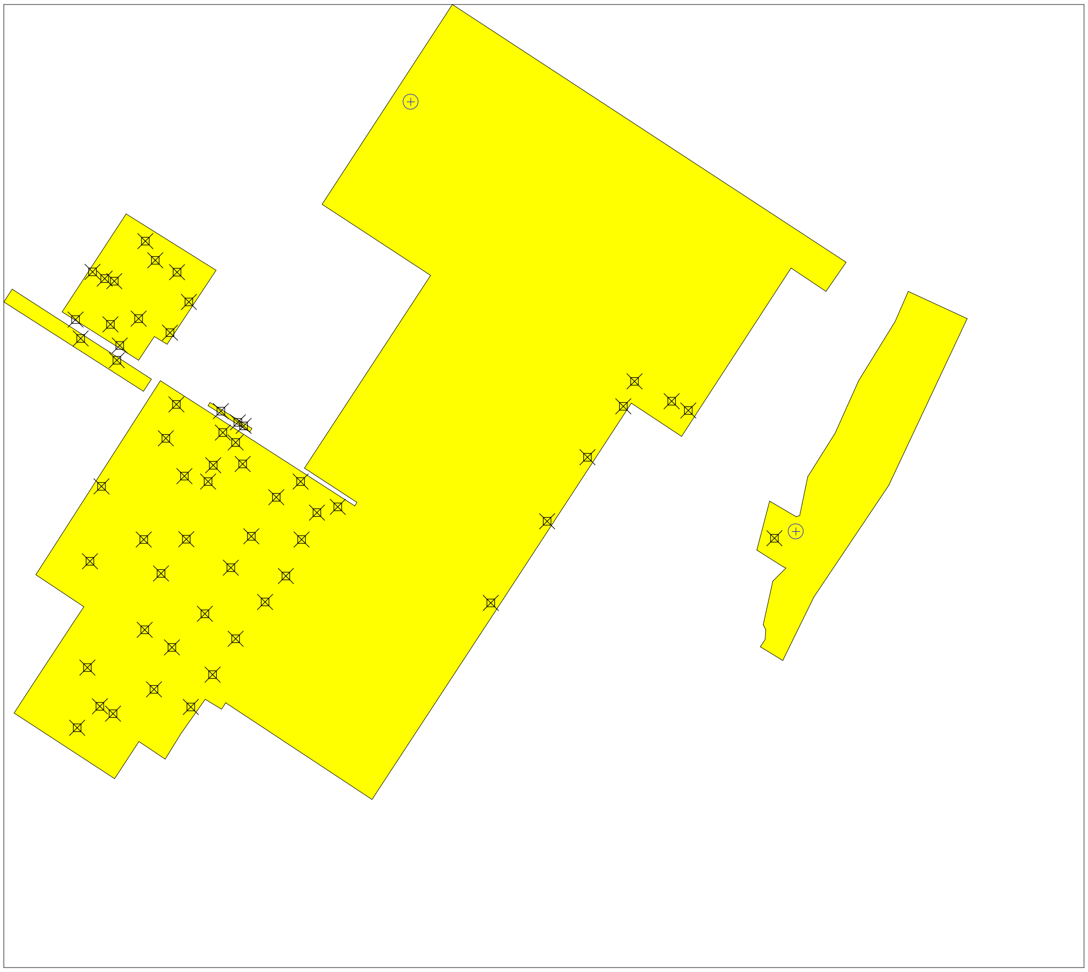
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679149.4920	3082933.0980	J-13S	0.077	Manual	T
679279.6830	3083075.4290	J-14S	0.077	Manual	T
679261.0980	3083016.3510	J-15S	0.07	Manual	T
679222.6340	3082840.1720	J-16S	0.072	Manual	T
679293.5600	3082950.4980	J-17S	0.079	Manual	T
679360.5700	3083026.4980	J-18S	0.074	Manual	T
679343.5810	3082969.5980	J-19S	0.077	Manual	T
679382.8640	3083009.1130	J-20S	0.83	Manual	T
679335.0020	3082941.1720	J-21S	0.073	Manual	T
679252.7130	3082781.0290	J-22S	0.072	Manual	T
679297.0010	3082840.6970	J-23S	0.071	Manual	T
679394.8070	3082971.8300	J-24S	0.073	Manual	T
679146.6460	3082549.7640	J-25S	0.064	Manual	T
679224.5850	3082683.1400	J-26S	0.072	Manual	T
679169.0760	3082537.3510	J-27S	0.072	Manual	T
679272.0040	3082652.6750	J-28S	0.073	Manual	T

679329.4380	3082711.0960	J-29S	0.071	Manual	T
679374.4420	3082791.3300	J-30S	0.074	Manual	T
679410.1490	3082845.8460	J-31S	0.079	Manual	T
679453.4760	3082914.1150	J-32S	0.076	Manual	T
679495.8840	3082940.9730	J-33S	0.073	Manual	T
679304.6530	3082548.6880	J-34S	0.073	Manual	T
679342.7410	3082605.3190	J-35S	0.076	Manual	T
679382.8900	3082667.5270	J-36S	0.074	Manual	T
679433.9450	3082731.6820	J-37S	0.071	Manual	T
679470.3570	3082776.7350	J-38S	0.078	Manual	T
679497.3310	3082840.3960	J-39S	0.074	Manual	T
679524.3310	3082886.8990	J-40S	0.074	Manual	T
679560.6070	3082897.2580	J-41S	0.072	Manual	T
679924.8150	3082872.3490	J-47S	0.076	Manual	T
679994.9690	3082983.5100	J-48S	0.076	Manual	T
680057.6580	3083072.0750	J-49S	0.079	Manual	T
680077.3540	3083115.5330	J-50S	0.076	Manual	T
679827.1150	3082729.7460	J-51S	0.074	Manual	T
680141.8730	3083080.8800	J-52S	0.076	Manual	T
680170.5600	3083064.6740	J-53S	0.075	Manual	T
679129.3320	3082802.5620	Composite 1	0.074	Manual	T
679124.7500	3082617.3010	Composite 3	0.076	Manual	T
679107.0750	3082512.5600	Composite 4	0.074	Manual	T
679240.6200	3082579.3320	Composite 2	0.075	Manual	T
679687.6145	3083602.0759	Composite 5	0.073	Random	

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.073	Manual	T
680357.8357	3082854.0284	J-62S	0.073	Random	

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.073	Manual	T
679396.8510	3083038.0640	J-64S	0.074	Manual	T
679386.3850	3083044.5490	J-63S	0.071	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.066	Manual	T
679113.1200	3083190.3150	J-66S	0.067	Manual	T

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.085	Manual	T
679104.2450	3083223.2620	J-02S	0.07	Manual	T
679155.0740	3083294.6960	J-03S	0.082	Manual	T
679171.2970	3083289.7960	J-04S	0.33	Manual	T
679225.8560	3083359.9740	J-05S	0.075	Manual	T
679164.8060	3083214.7100	J-06S	0.07	Manual	T
679242.7260	3083326.5280	J-07S	0.07	Manual	T
679181.2750	3083178.2880	J-08S	0.071	Manual	T
679213.7730	3083224.9730	J-09S	0.072	Manual	T
679280.5440	3083305.6810	J-10S	0.073	Manual	T
679268.7700	3083200.3260	J-11S	0.083	Manual	T
679301.1600	3083254.0340	J-12S	0.072	Manual	T

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,

β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

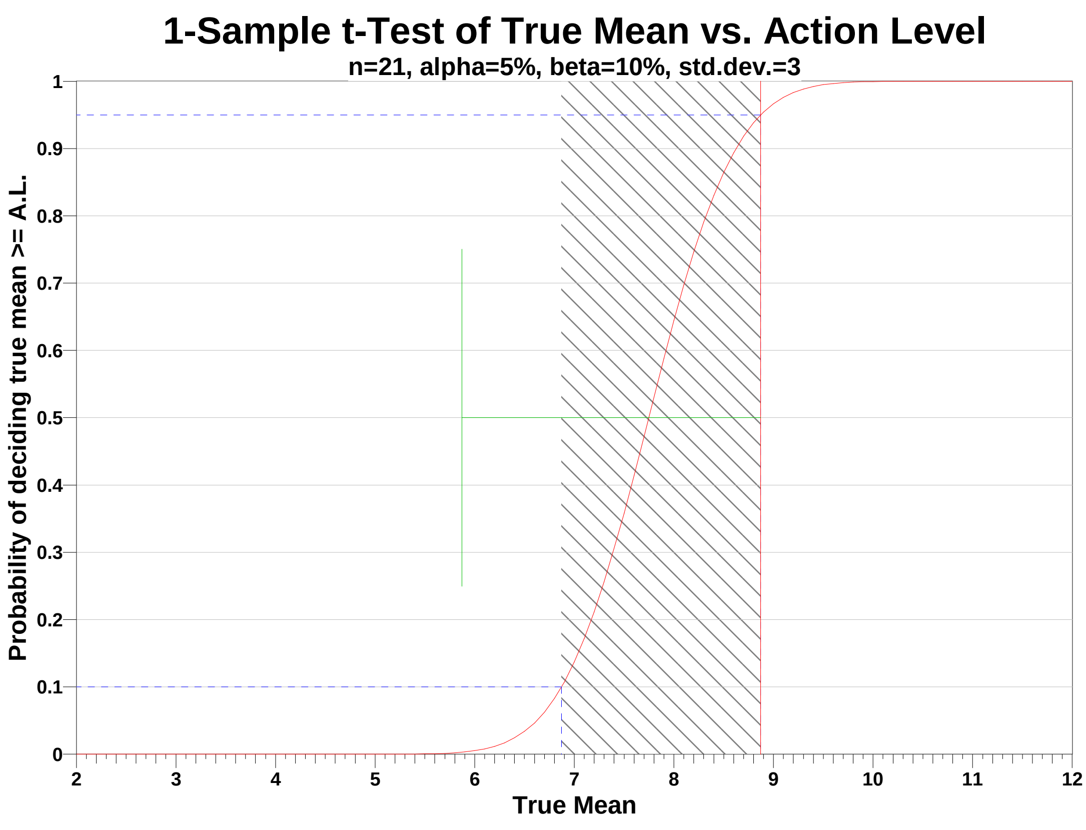
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,

3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples								
AL=8.87154		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	497	126	393	99	330	83
	β=10	394	100	302	76	247	62
	β=15	331	84	247	63	198	50
LBGR=80	β=5	126	33	99	26	83	22
	β=10	100	26	76	20	62	16
	β=15	84	22	63	17	50	13
LBGR=70	β=5	57	16	45	12	38	10
	β=10	45	13	35	10	28	8
	β=15	38	11	29	8	23	6

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling

The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10
0	0	0.064	0.066	0.067	0.07	0.07	0.07	0.07	0.07	0.071

10	0.071	0.071	0.071	0.071	0.072	0.072	0.072	0.072	0.072	0.072
20	0.072	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073
30	0.073	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.074
40	0.075	0.075	0.075	0.076	0.076	0.076	0.076	0.076	0.076	0.076
50	0.077	0.077	0.077	0.078	0.079	0.079	0.079	0.082	0.083	0.085
60	0.085	0.33	0.83							

SUMMARY STATISTICS								
n	63							
Min	0							
Max	0.83							
Range	0.83							
Mean	0.088889							
Median	0.074							
Variance	0.010153							
StdDev	0.10076							
Std Error	0.012695							
Skewness	6.8407							
Interquartile Range	0.004							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.0662	0.07	0.072	0.074	0.076	0.0808	0.085	0.83

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	7.355	3.218	Yes

The test statistic 7.355 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	0.83

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the

suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.3953
Lilliefors 5% Critical Value	0.1125

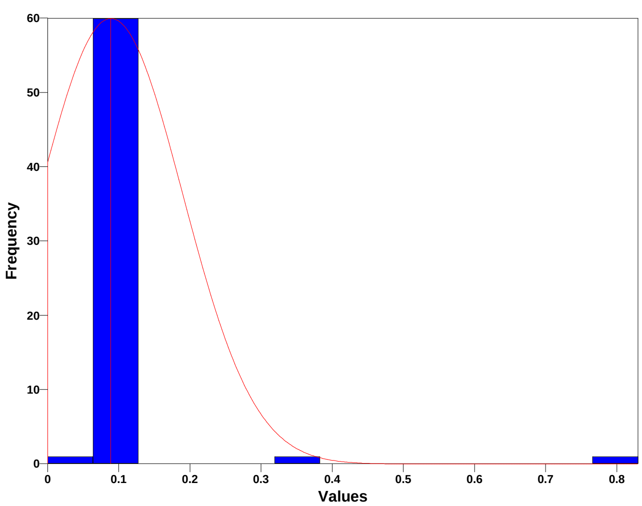
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

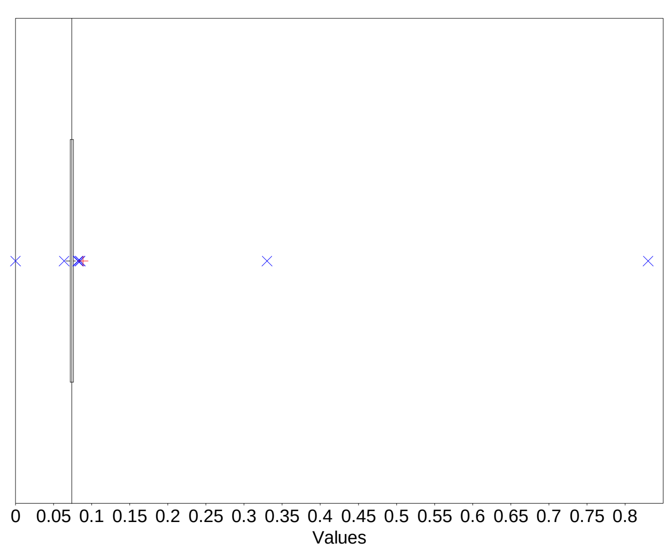
Data Plots
Graphical displays of the data are shown below.

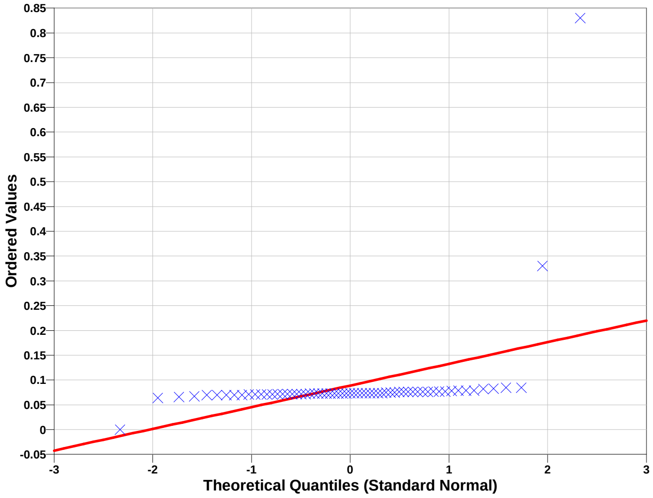
The Histogram is a plot of the fraction of the n observed data that fall within specified data "bins." A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.4836
Lilliefors 5% Critical Value	0.1116

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.1101
95% Non-Parametric (Chebyshev) UCL	0.1442

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.1442) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=63 data,
AL is the action level or threshold (8.87154),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=62 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-691.82	1.6698	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
63	39	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

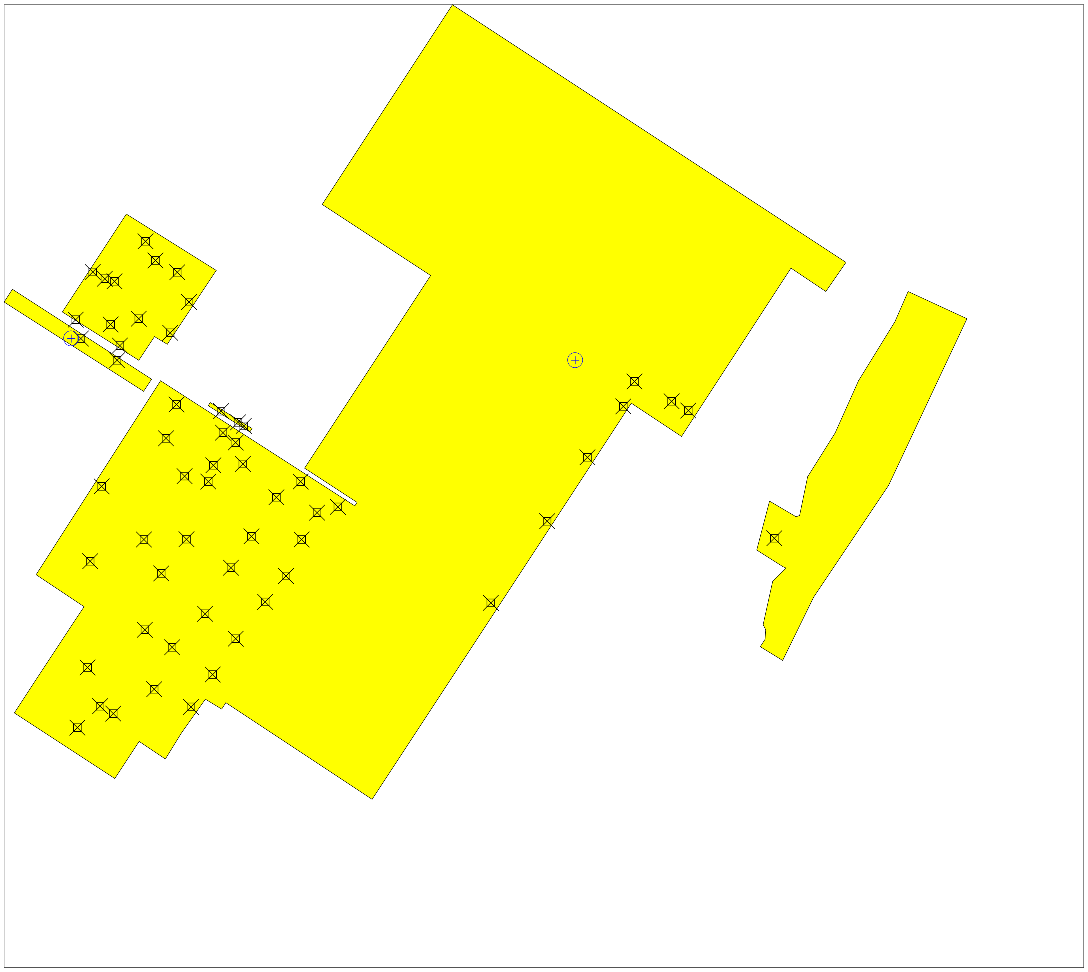
SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.

[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.

[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.

[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679149.4920	3082933.0980	J-13S	0.077	Manual	T
679279.6830	3083075.4290	J-14S	0.077	Manual	T
679261.0980	3083016.3510	J-15S	0.07	Manual	T
679222.6340	3082840.1720	J-16S	0.072	Manual	T
679293.5600	3082950.4980	J-17S	0.079	Manual	T
679360.5700	3083026.4980	J-18S	0.074	Manual	T
679343.5810	3082969.5980	J-19S	0.077	Manual	T
679382.8640	3083009.1130	J-20S	0.83	Manual	T
679335.0020	3082941.1720	J-21S	0.073	Manual	T
679252.7130	3082781.0290	J-22S	0.072	Manual	T
679297.0010	3082840.6970	J-23S	0.071	Manual	T
679394.8070	3082971.8300	J-24S	0.073	Manual	T
679146.6460	3082549.7640	J-25S	0.064	Manual	T
679224.5850	3082683.1400	J-26S	0.072	Manual	T
679169.0760	3082537.3510	J-27S	0.072	Manual	T
679272.0040	3082652.6750	J-28S	0.073	Manual	T

679329.4380	3082711.0960	J-29S	0.071	Manual	T
679374.4420	3082791.3300	J-30S	0.074	Manual	T
679410.1490	3082845.8460	J-31S	0.079	Manual	T
679453.4760	3082914.1150	J-32S	0.076	Manual	T
679495.8840	3082940.9730	J-33S	0.073	Manual	T
679304.6530	3082548.6880	J-34S	0.073	Manual	T
679342.7410	3082605.3190	J-35S	0.076	Manual	T
679382.8900	3082667.5270	J-36S	0.074	Manual	T
679433.9450	3082731.6820	J-37S	0.071	Manual	T
679470.3570	3082776.7350	J-38S	0.078	Manual	T
679497.3310	3082840.3960	J-39S	0.074	Manual	T
679524.3310	3082886.8990	J-40S	0.074	Manual	T
679560.6070	3082897.2580	J-41S	0.072	Manual	T
679924.8150	3082872.3490	J-47S	0.076	Manual	T
679994.9690	3082983.5100	J-48S	0.076	Manual	T
680057.6580	3083072.0750	J-49S	0.079	Manual	T
680077.3540	3083115.5330	J-50S	0.076	Manual	T
679827.1150	3082729.7460	J-51S	0.074	Manual	T
680141.8730	3083080.8800	J-52S	0.076	Manual	T
680170.5600	3083064.6740	J-53S	0.075	Manual	T
679129.3320	3082802.5620	Composite 1	0.074	Manual	T
679124.7500	3082617.3010	Composite 3	0.076	Manual	T
679107.0750	3082512.5600	Composite 4	0.074	Manual	T
679240.6200	3082579.3320	Composite 2	0.075	Manual	T
679974.2022	3083152.2432	Composite 5	0.073	Random	

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	0.073	Manual	T

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	0.073	Manual	T
679396.8510	3083038.0640	J-64S	0.074	Manual	T
679386.3850	3083044.5490	J-63S	0.071	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.066	Manual	T
679113.1200	3083190.3150	J-66S	0.067	Manual	T
679096.4233	3083190.2180	J-01S	0.085	Random	

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.085	Manual	T
679104.2450	3083223.2620	J-02S	0.07	Manual	T
679155.0740	3083294.6960	J-03S	0.082	Manual	T
679171.2970	3083289.7960	J-04S	0.33	Manual	T
679225.8560	3083359.9740	J-05S	0.075	Manual	T
679164.8060	3083214.7100	J-06S	0.07	Manual	T
679242.7260	3083326.5280	J-07S	0.07	Manual	T
679181.2750	3083178.2880	J-08S	0.071	Manual	T
679213.7730	3083224.9730	J-09S	0.072	Manual	T
679280.5440	3083305.6810	J-10S	0.073	Manual	T
679268.7700	3083200.3260	J-11S	0.083	Manual	T
679301.1600	3083254.0340	J-12S	0.072	Manual	T

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,

β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

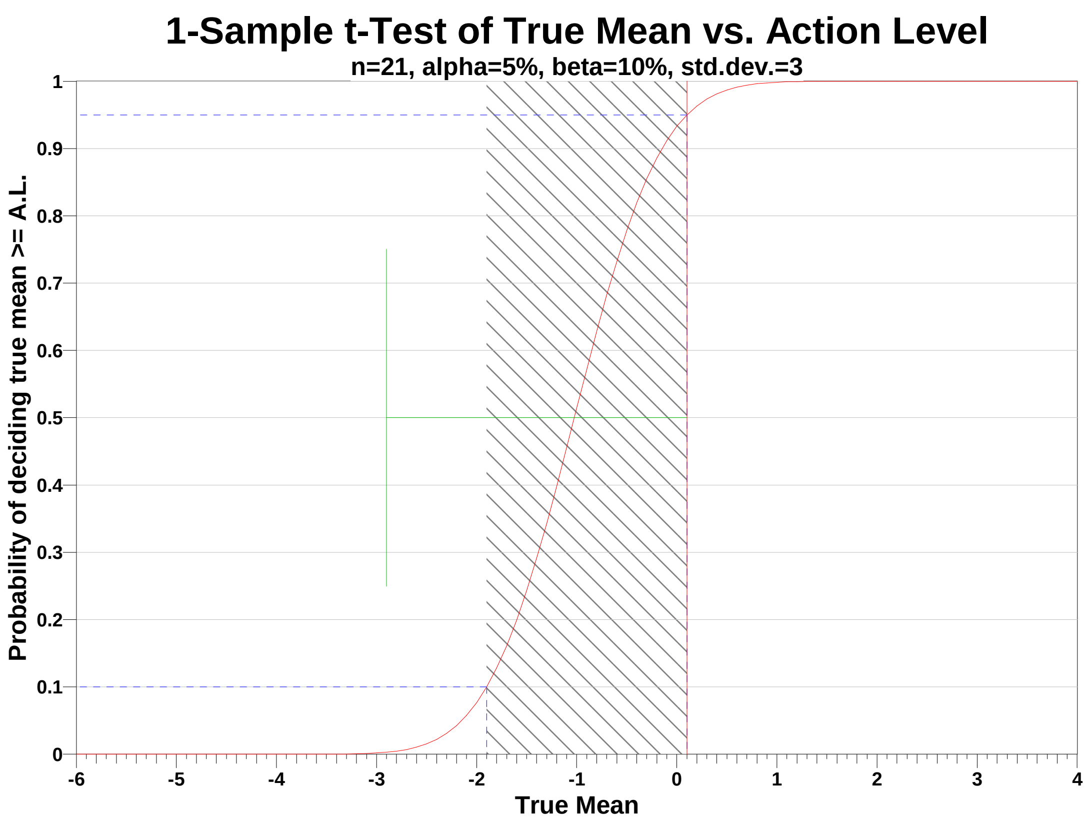
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,

3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.1		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	3895984	973997	3082986	770748	2588149	647038
	β=10	3082987	770748	2365020	591256	1934301	483576
	β=15	2588150	647039	1934301	483576	1546841	386711
LBGR=80	β=5	973997	243501	770748	192688	647038	161760
	β=10	770748	192688	591256	147815	483576	120895
	β=15	647039	161761	483576	120895	386711	96679
LBGR=70	β=5	432889	108224	342555	85640	287573	71894
	β=10	342556	85640	262781	65696	214923	53732
	β=15	287574	71895	214923	53732	171872	42969

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling

The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

The following data points were entered by the user for analysis.

Rank	1	2	3	4	5	6	7	8	9	10
0	0.064	0.066	0.067	0.07	0.07	0.07	0.07	0.07	0.071	0.071

10	0.071	0.071	0.071	0.072	0.072	0.072	0.072	0.072	0.072	0.072
20	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073
30	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.074	0.075
40	0.075	0.075	0.076	0.076	0.076	0.076	0.076	0.076	0.076	0.077
50	0.077	0.077	0.078	0.079	0.079	0.079	0.082	0.083	0.085	0.085
60	0.33	0.83								

SUMMARY STATISTICS	
n	62
Min	0.064
Max	0.83
Range	0.766
Mean	0.090323
Median	0.074
Variance	0.010188
StdDev	0.10094
Std Error	0.012819
Skewness	6.8884
Interquartile Range	0.004

Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.064	0.06745	0.07	0.072	0.074	0.076	0.0811	0.085	0.83

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	7.328	3.212	Yes

The test statistic 7.328 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	0.83

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the

suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.4083
Lilliefors 5% Critical Value	0.1134

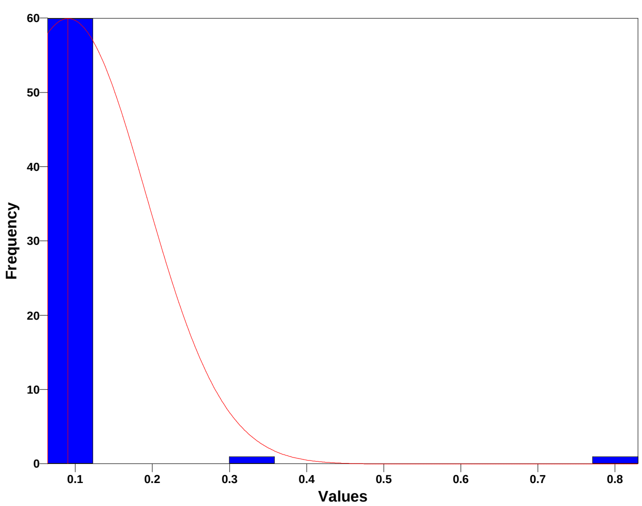
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

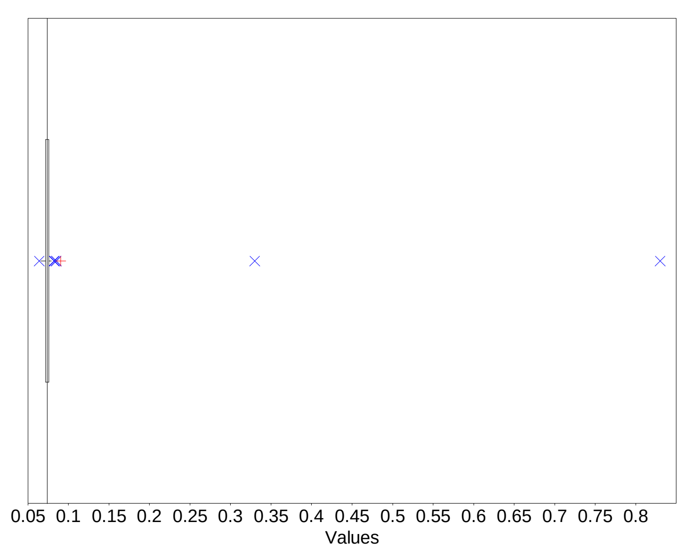
Data Plots
Graphical displays of the data are shown below.

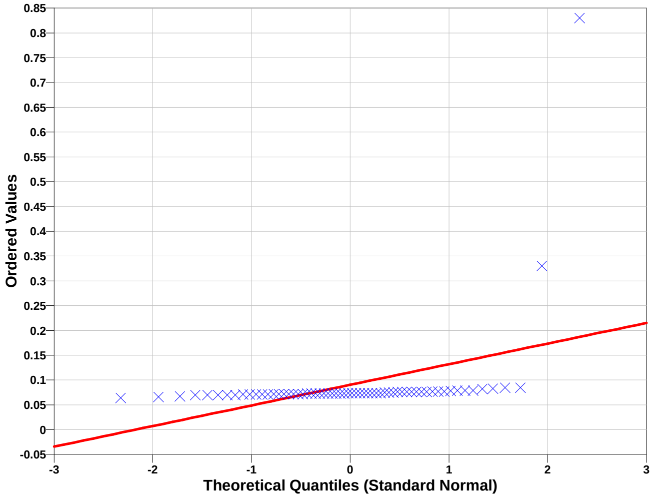
The Histogram is a plot of the fraction of the n observed data that fall within specified data "bins." A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25th and 75th percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The pth quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.4888
Lilliefors 5% Critical Value	0.1125

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.1117
95% Non-Parametric (Chebyshev) UCL	0.1462

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.1462) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=62 data,
AL is the action level or threshold (0.1),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=61 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-0.75494	1.6702	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
60	38	Reject

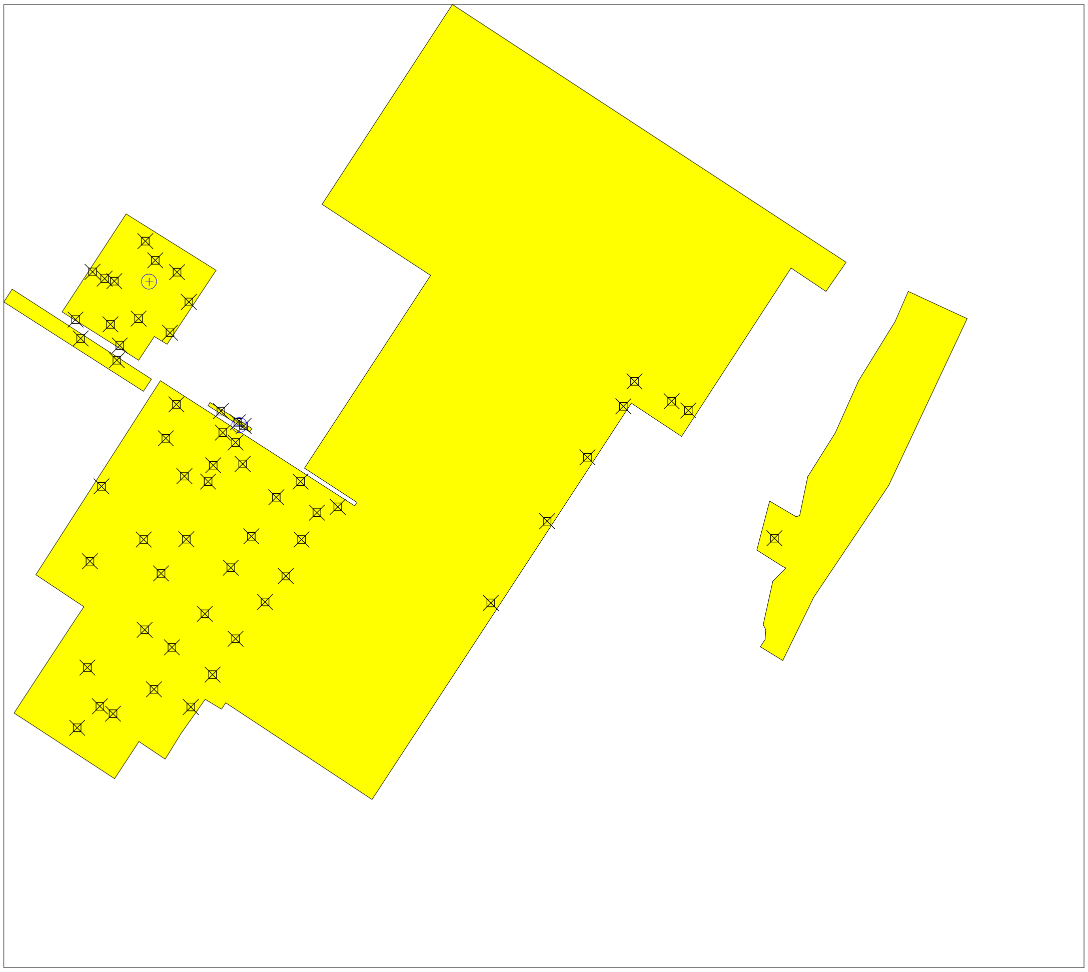
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679149.4920	3082933.0980	J-13S	4.9	Manual	T
679279.6830	3083075.4290	J-14S	0.6	Manual	T
679261.0980	3083016.3510	J-15S	3.4	Manual	T
679222.6340	3082840.1720	J-16S	3.2	Manual	T
679293.5600	3082950.4980	J-17S	2.4	Manual	T
679360.5700	3083026.4980	J-18S	3.2	Manual	T
679343.5810	3082969.5980	J-19S	15	Manual	T
679382.8640	3083009.1130	J-20S	2.5	Manual	T
679335.0020	3082941.1720	J-21S	0.59	Manual	T
679252.7130	3082781.0290	J-22S	4.1	Manual	T
679297.0010	3082840.6970	J-23S	3.5	Manual	T
679394.8070	3082971.8300	J-24S	1.2	Manual	T
679146.6460	3082549.7640	J-25S	7.4	Manual	T
679224.5850	3082683.1400	J-26S	1.7	Manual	T
679169.0760	3082537.3510	J-27S	3.7	Manual	T
679272.0040	3082652.6750	J-28S	6.1	Manual	T

679329.4380	3082711.0960	J-29S	2.7	Manual	T
679374.4420	3082791.3300	J-30S	2.2	Manual	T
679410.1490	3082845.8460	J-31S	0.76	Manual	T
679453.4760	3082914.1150	J-32S	2.9	Manual	T
679495.8840	3082940.9730	J-33S	1.9	Manual	T
679304.6530	3082548.6880	J-34S	1.5	Manual	T
679342.7410	3082605.3190	J-35S	1.8	Manual	T
679382.8900	3082667.5270	J-36S	6	Manual	T
679433.9450	3082731.6820	J-37S	8.8	Manual	T
679470.3570	3082776.7350	J-38S	2.5	Manual	T
679497.3310	3082840.3960	J-39S	1.4	Manual	T
679524.3310	3082886.8990	J-40S	2.3	Manual	T
679560.6070	3082897.2580	J-41S	2.2	Manual	T
679924.8150	3082872.3490	J-47S	2.7	Manual	T
679994.9690	3082983.5100	J-48S	3.4	Manual	T
680057.6580	3083072.0750	J-49S	3.6	Manual	T
680077.3540	3083115.5330	J-50S	3.2	Manual	T
679827.1150	3082729.7460	J-51S	2.2	Manual	T
680141.8730	3083080.8800	J-52S	4	Manual	T
680170.5600	3083064.6740	J-53S	3.9	Manual	T
679129.3320	3082802.5620	Composite 1	1.6	Manual	T
679240.6200	3082579.3320	Composite 2	3	Manual	T
679124.7500	3082617.3010	Composite 3	2.1	Manual	T
679107.0750	3082512.5600	Composite 4	1.6	Manual	T

Area: Area 2

X Coord	Y Coord	Label	Value	Type	Historical
680320.6560	3082842.6400	Composite 5	4.4	Manual	T

Area: Area 3

X Coord	Y Coord	Label	Value	Type	Historical
679356.9310	3083064.0350	J-62S	2.1	Manual	T
679386.3850	3083044.5490	J-63S	4	Manual	T
679396.8510	3083038.0640	J-64S	2.3	Manual	T
679389.4539	3083039.3751	J-65S	0.9	Random	

Area: Area 4

X Coord	Y Coord	Label	Value	Type	Historical
679175.7550	3083152.6270	J-65S	0.9	Manual	T
679113.1200	3083190.3150	J-66S	0.53	Manual	T

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.63	Manual	T
679104.2450	3083223.2620	J-02S	3.9	Manual	T
679155.0740	3083294.6960	J-03S	1.2	Manual	T
679171.2970	3083289.7960	J-04S	2.1	Manual	T
679225.8560	3083359.9740	J-05S	0.8	Manual	T
679164.8060	3083214.7100	J-06S	3.7	Manual	T
679242.7260	3083326.5280	J-07S	4.3	Manual	T
679181.2750	3083178.2880	J-08S	5.2	Manual	T
679213.7730	3083224.9730	J-09S	1.8	Manual	T
679280.5440	3083305.6810	J-10S	1	Manual	T
679268.7700	3083200.3260	J-11S	0.98	Manual	T
679301.1600	3083254.0340	J-12S	0.58	Manual	T
679232.5619	3083288.7154		0	Random	

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability (1-β) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,

Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

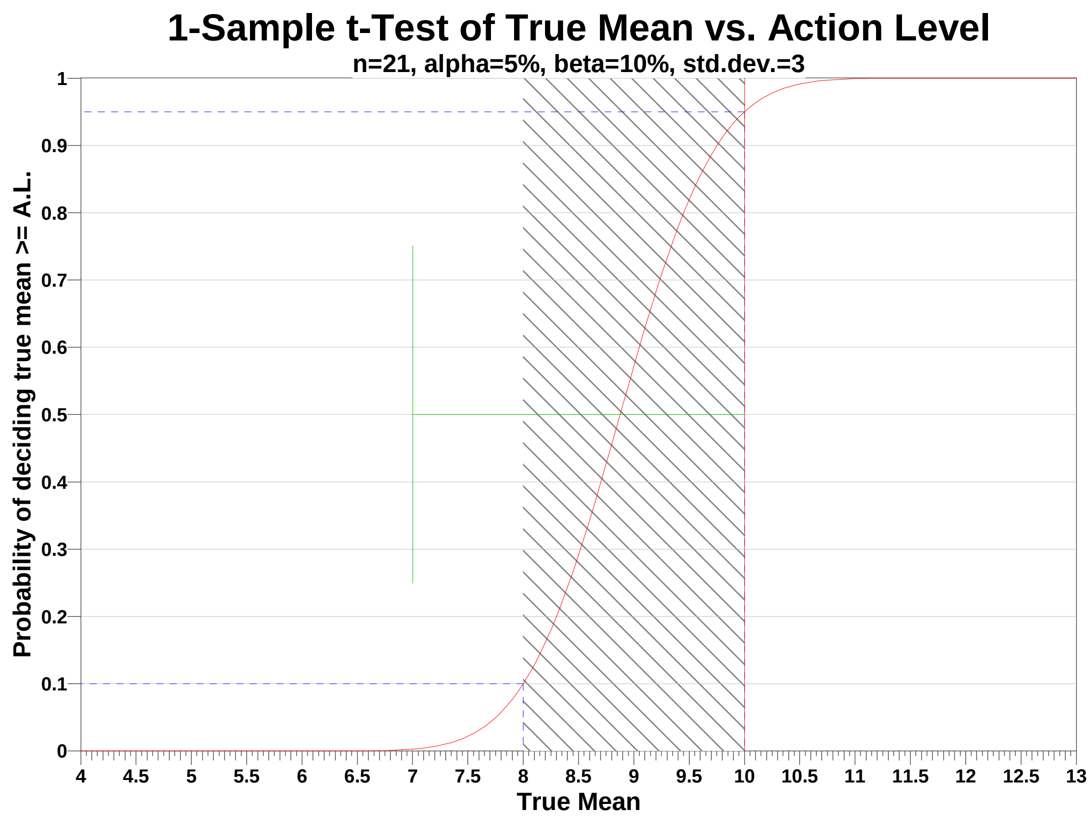
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30

or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples								
AL=10		α=5		α=10		α=15		
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	391	99	310	78	260	66
	β=10	310	79	238	60	194	49
	β=15	261	67	195	50	156	40
LBGR=80	β=5	99	26	78	21	66	17
	β=10	79	21	60	16	49	13
	β=15	67	18	50	13	40	11
LBGR=70	β=5	45	13	36	10	30	8
	β=10	36	10	28	8	23	6
	β=15	31	9	23	7	18	5

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis
The following data points were entered by the user for analysis.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**

0	0	0.53	0.58	0.59	0.6	0.63	0.76	0.8	0.9	0.9
10	0.98	1	1.2	1.2	1.4	1.5	1.6	1.6	1.7	1.8
20	1.8	1.9	2.1	2.1	2.1	2.1	2.2	2.2	2.2	2.3
30	2.3	2.4	2.5	2.5	2.7	2.7	2.9	3	3.2	3.2
40	3.2	3.4	3.4	3.5	3.6	3.7	3.7	3.9	3.9	4
50	4	4.1	4.3	4.4	4.4	4.9	5.2	6	6.1	7.4
60	8.8	15								

SUMMARY STATISTICS								
n	62							
Min	0							
Max	15							
Range	15							
Mean	2.8963							
Median	2.35							
Variance	5.3925							
StdDev	2.3222							
Std Error	0.29492							
Skewness	2.7155							
Interquartile Range	2.275							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.5815	0.669	1.475	2.35	3.75	5.11	7.205	15

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	5.193	3.2	Yes

The test statistic 5.193 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	15

Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.1061
Lilliefors 5% Critical Value	0.1153

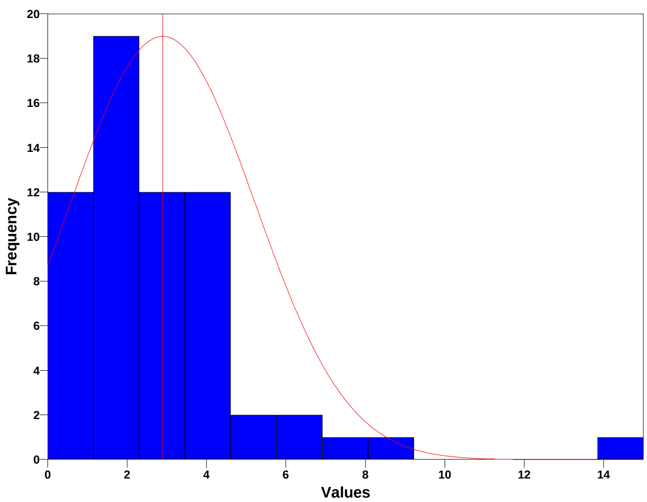
The calculated Lilliefors test statistic is less than the 5% Lilliefors critical value, so the test cannot reject the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do appear to follow a normal distribution at the 5% level of significance.

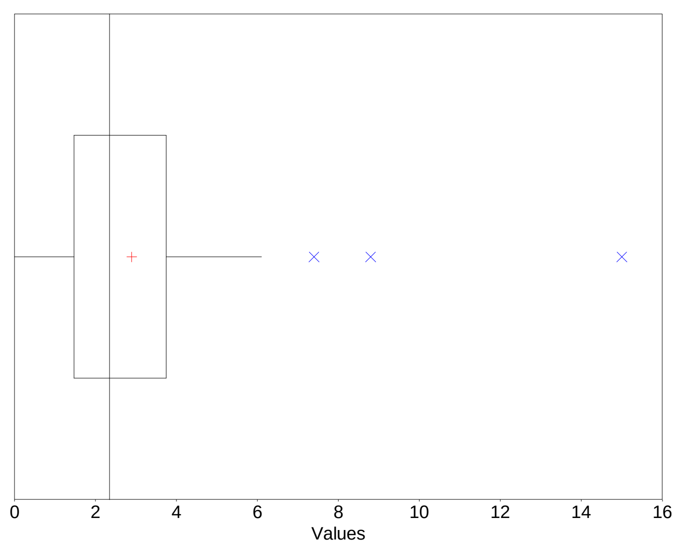
Data Plots
Graphical displays of the data are shown below.

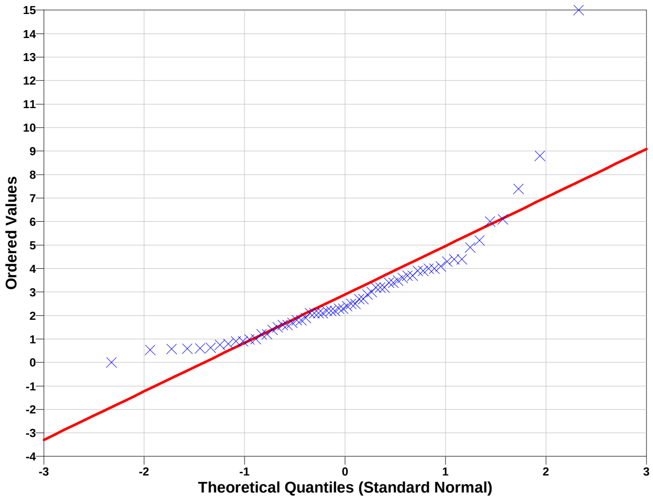
The Histogram is a plot of the fraction of the n observed data that fall within specified data "bins." A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.1457
Lilliefors 5% Critical Value	0.1125

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	3.389
95% Non-Parametric (Chebyshev) UCL	4.182

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (4.182) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=62 data,
AL is the action level or threshold (10),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=61 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-24.087	1.6702	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
61	38	Reject

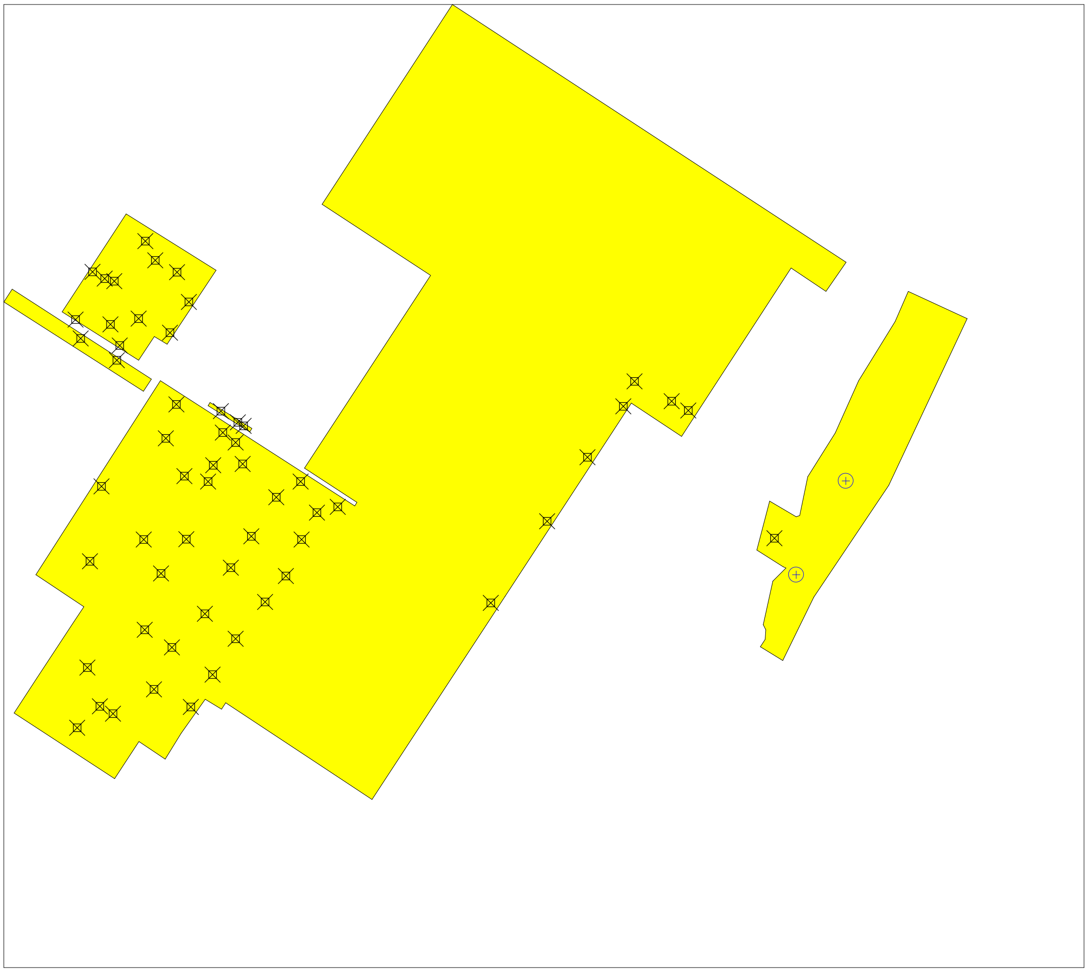
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	60
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Area: Area 1					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679149.4920	3082933.0980	J-13S	4.9	Manual	T
679279.6830	3083075.4290	J-14S	0.6	Manual	T
679261.0980	3083016.3510	J-15S	3.4	Manual	T
679222.6340	3082840.1720	J-16S	3.2	Manual	T
679293.5600	3082950.4980	J-17S	2.4	Manual	T
679360.5700	3083026.4980	J-18S	3.2	Manual	T
679343.5810	3082969.5980	J-19S	15	Manual	T
679382.8640	3083009.1130	J-20S	2.5	Manual	T
679335.0020	3082941.1720	J-21S	0.59	Manual	T
679252.7130	3082781.0290	J-22S	4.1	Manual	T
679297.0010	3082840.6970	J-23S	3.5	Manual	T
679394.8070	3082971.8300	J-24S	1.2	Manual	T
679146.6460	3082549.7640	J-25S	7.4	Manual	T
679224.5850	3082683.1400	J-26S	1.7	Manual	T
679169.0760	3082537.3510	J-27S	3.7	Manual	T
679272.0040	3082652.6750	J-28S	6.1	Manual	T

679329.4380	3082711.0960	J-29S	2.7	Manual	T
679374.4420	3082791.3300	J-30S	2.2	Manual	T
679410.1490	3082845.8460	J-31S	0.76	Manual	T
679453.4760	3082914.1150	J-32S	2.9	Manual	T
679495.8840	3082940.9730	J-33S	1.9	Manual	T
679304.6530	3082548.6880	J-34S	1.5	Manual	T
679342.7410	3082605.3190	J-35S	1.8	Manual	T
679382.8900	3082667.5270	J-36S	6	Manual	T
679433.9450	3082731.6820	J-37S	8.8	Manual	T
679470.3570	3082776.7350	J-38S	2.5	Manual	T
679497.3310	3082840.3960	J-39S	1.4	Manual	T
679524.3310	3082886.8990	J-40S	2.3	Manual	T
679560.6070	3082897.2580	J-41S	2.2	Manual	T
679924.8150	3082872.3490	J-47S	2.7	Manual	T
679994.9690	3082983.5100	J-48S	3.4	Manual	T
680057.6580	3083072.0750	J-49S	3.6	Manual	T
680077.3540	3083115.5330	J-50S	3.2	Manual	T
679827.1150	3082729.7460	J-51S	2.2	Manual	T
680141.8730	3083080.8800	J-52S	4	Manual	T
680170.5600	3083064.6740	J-53S	3.9	Manual	T
679129.3320	3082802.5620	Composite 1	1.6	Manual	T
679240.6200	3082579.3320	Composite 2	3	Manual	T
679124.7500	3082617.3010	Composite 3	2.1	Manual	T
679107.0750	3082512.5600	Composite 4	1.6	Manual	T

Area: Area 2					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
680320.6560	3082842.6400	Composite 5	4.4	Manual	T
680445.0613	3082942.1178	J-62S	2.1	Random	
680358.3341	3082778.8501	J-63S	4	Random	

Area: Area 3					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679356.9310	3083064.0350	J-62S	2.1	Manual	T
679386.3850	3083044.5490	J-63S	4	Manual	T
679396.8510	3083038.0640	J-64S	2.3	Manual	T

Area: Area 4					
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**
679175.7550	3083152.6270	J-65S	0.9	Manual	T
679113.1200	3083190.3150	J-66S	0.53	Manual	T

Area: Area 5

X Coord	Y Coord	Label	Value	Type	Historical
679133.4290	3083306.3130	J-01S	0.63	Manual	T
679104.2450	3083223.2620	J-02S	3.9	Manual	T
679155.0740	3083294.6960	J-03S	1.2	Manual	T
679171.2970	3083289.7960	J-04S	2.1	Manual	T
679225.8560	3083359.9740	J-05S	0.8	Manual	T
679164.8060	3083214.7100	J-06S	3.7	Manual	T
679242.7260	3083326.5280	J-07S	4.3	Manual	T
679181.2750	3083178.2880	J-08S	5.2	Manual	T
679213.7730	3083224.9730	J-09S	1.8	Manual	T
679280.5440	3083305.6810	J-10S	1	Manual	T
679268.7700	3083200.3260	J-11S	0.98	Manual	T
679301.1600	3083254.0340	J-12S	0.58	Manual	T

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,

β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

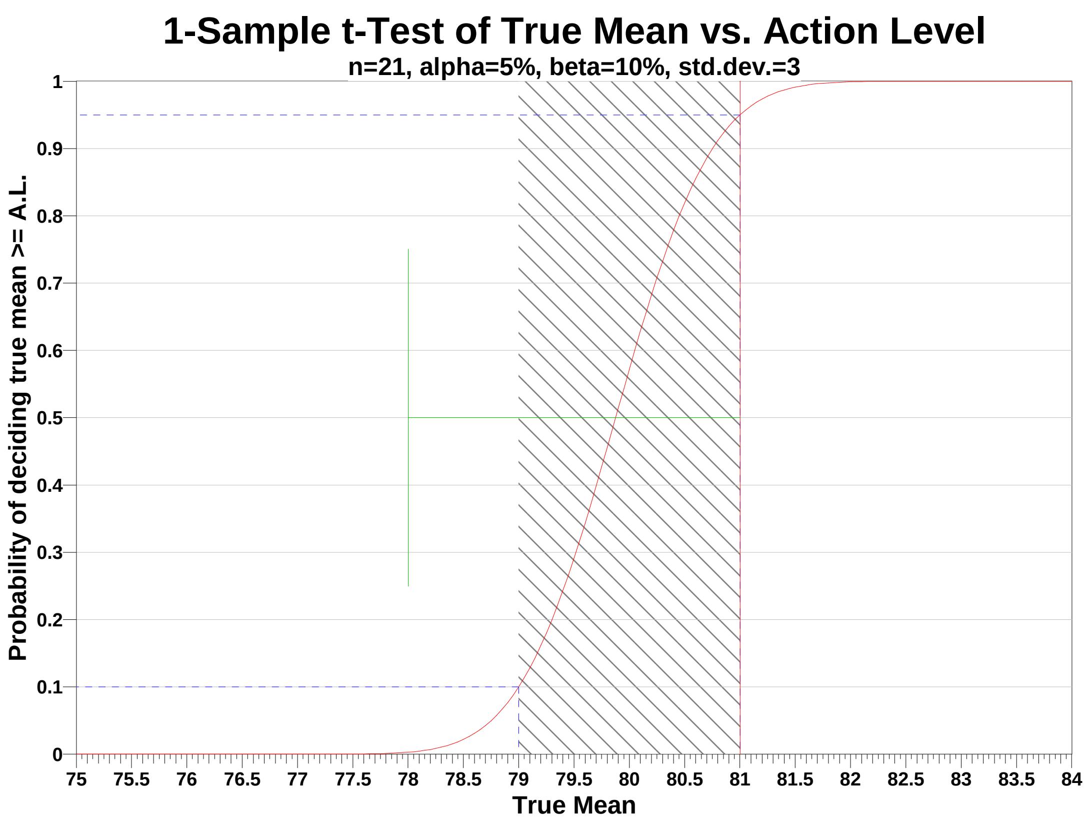
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,

3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples								
AL=81		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	8	3	6	2	5	2
	β=10	7	3	5	2	4	2
	β=15	6	3	4	2	3	2
LBGR=80	**β=5**	3	2	2	2	2	1
	β=10	3	2	2	2	2	1
	β=15	3	2	2	2	2	1
LBGR=70	**β=5**	3	2	2	1	1	1
	β=10	2	2	2	1	1	1
	β=15	2	2	2	1	1	1

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis
The following data points were entered by the user for analysis.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
0	0	0.53	0.58	0.59	0.6	0.63	0.63	0.76	0.8	0.9

10	0.9	0.98	1	1.2	1.2	1.4	1.5	1.6	1.6	1.7
20	1.8	1.8	1.9	2.1	2.1	2.1	2.1	2.2	2.2	2.2
30	2.3	2.3	2.4	2.5	2.5	2.7	2.7	2.9	3	3.2
40	3.2	3.2	3.4	3.4	3.5	3.6	3.7	3.7	3.9	3.9
50	4	4	4	4.1	4.3	4.4	4.4	4.9	5.2	6
60	6.1	7.4	8.8	15						

SUMMARY STATISTICS								
n	64							
Min	0							
Max	15							
Range	15							
Mean	2.8781							
Median	2.35							
Variance	5.3219							
StdDev	2.3069							
Std Error	0.28836							
Skewness	2.6897							
Interquartile Range	2.425							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.5825	0.63	1.425	2.35	3.85	5.05	7.075	15

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	5.23	3.218	Yes

The test statistic 5.23 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	15

Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.1022
Lilliefors 5% Critical Value	0.1125

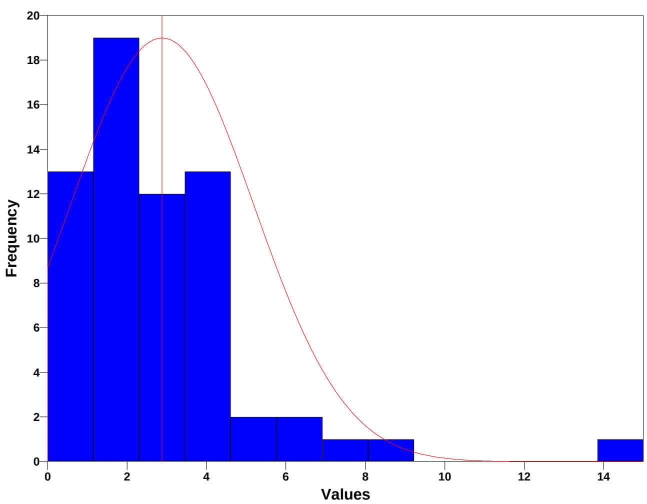
The calculated Lilliefors test statistic is less than the 5% Lilliefors critical value, so the test cannot reject the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do appear to follow a normal distribution at the 5% level of significance.

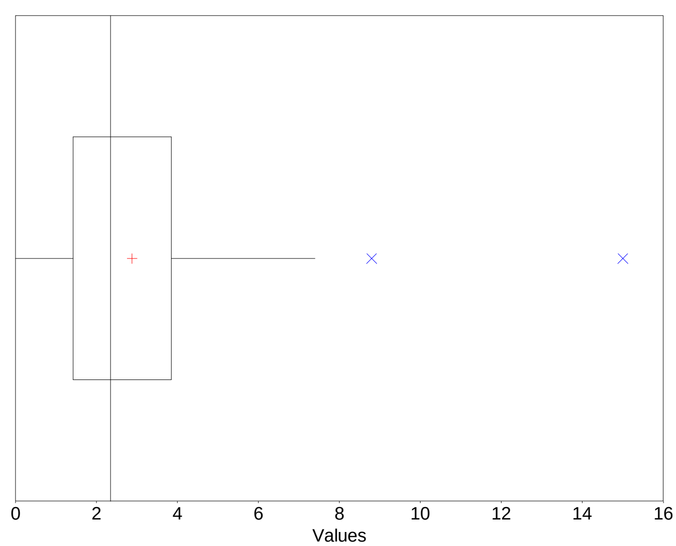
Data Plots
Graphical displays of the data are shown below.

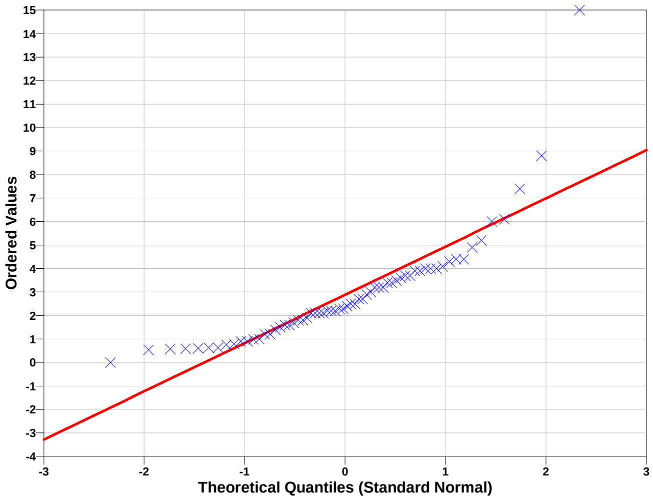
The Histogram is a plot of the fraction of the n observed data that fall within specified data "bins." A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_n, for which a fraction p of the distribution is less than x_n. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.1453
Lilliefors 5% Critical Value	0.1108

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	3.36
95% Non-Parametric (Chebyshev) UCL	4.135

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (4.135) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=64 data,
AL is the action level or threshold (81),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=63 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-270.91	1.6694	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
64	39	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

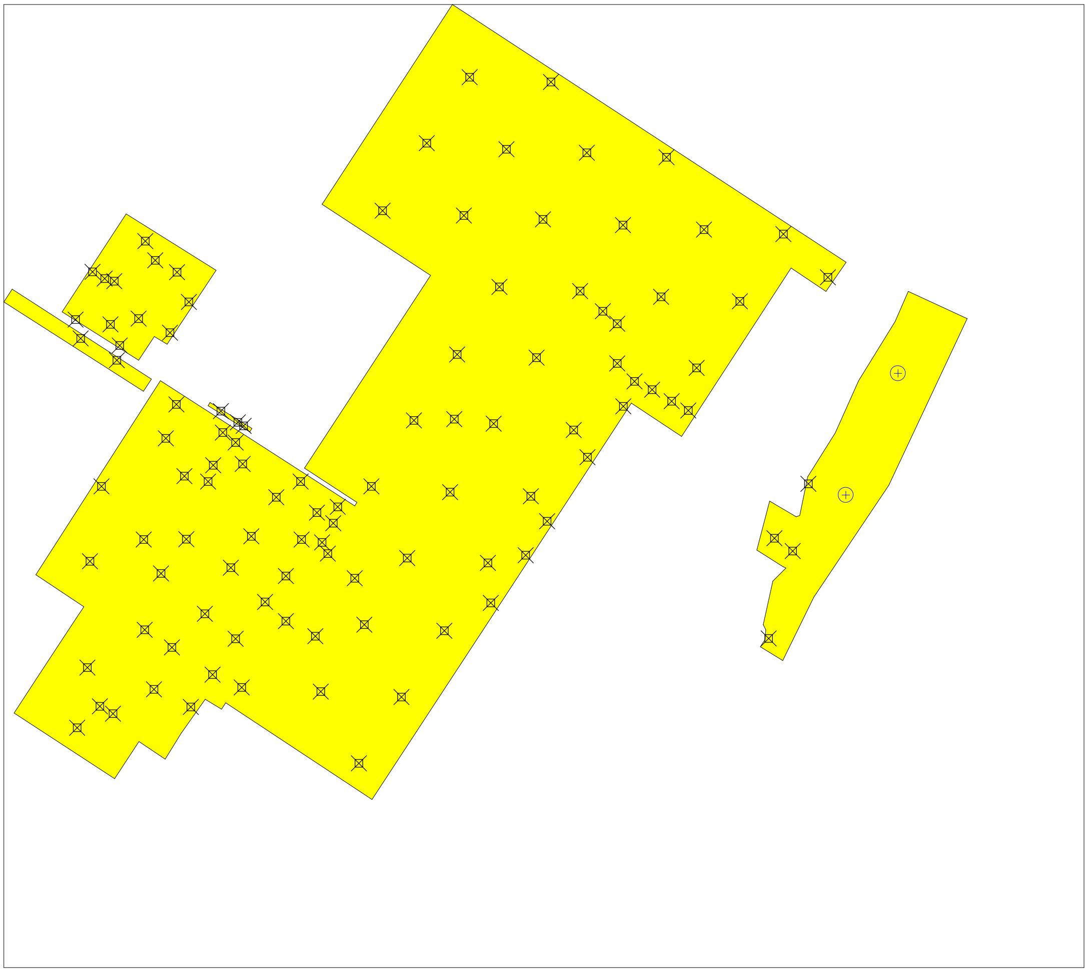
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

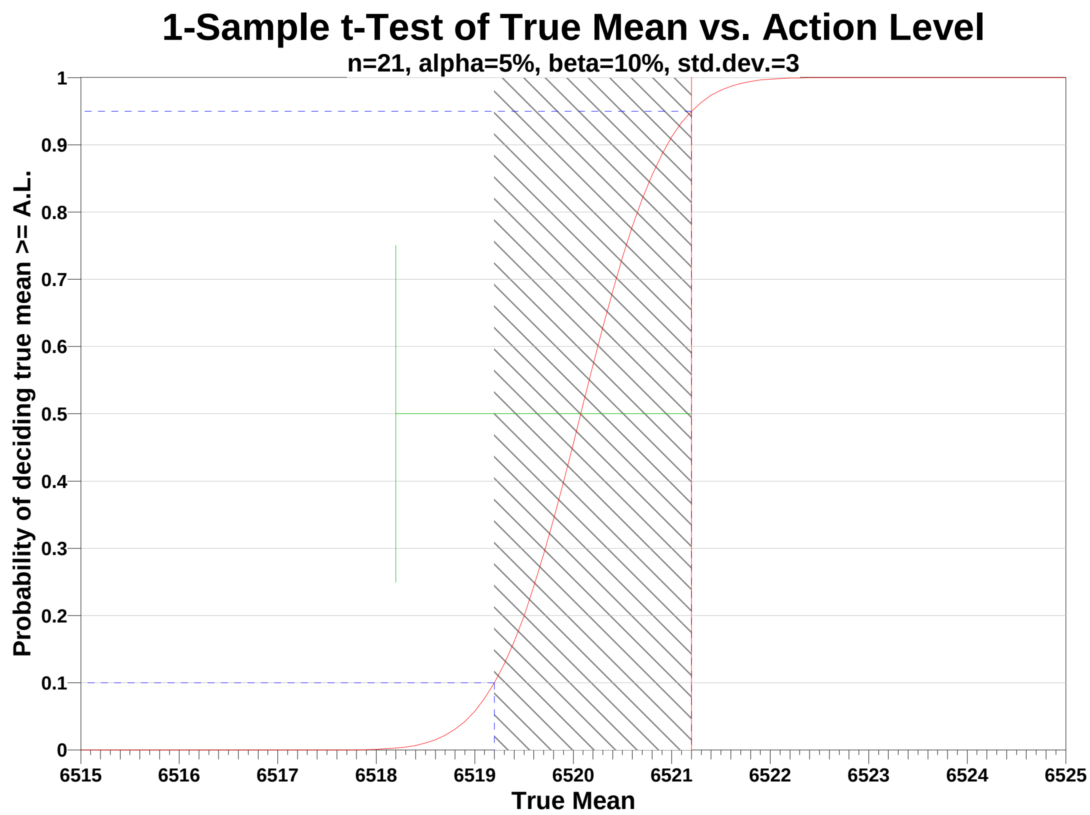
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=6521.2		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	2	2	1	1	1	1
	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1
LBGR=80	β=5	2	2	1	1	1	1
	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1
LBGR=70	β=5	2	2	1	1	1	1

	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS	
n	112
Min	0
Max	35900
Range	35900
Mean	5946.6
Median	4265
Variance	3.8705e+007
StdDev	6221.3
Std Error	587.86
Skewness	2.5682
Interquartile Range	5567.5

Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
79.17	808.1	1128	2073	4265	7640	1.389e+004	1.679e+004	3.574e+004

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	4.804	3.411	Yes

The test statistic 4.804 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	35900

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.22
Lilliefors 5% Critical Value	0.08448

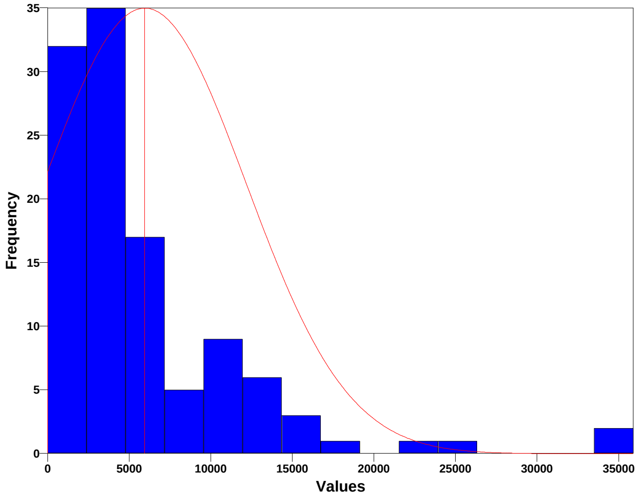
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

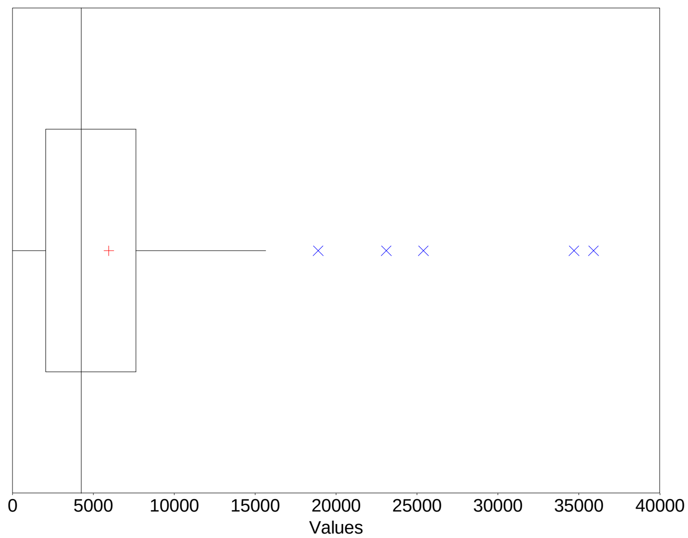
Data Plots
Graphical displays of the data are shown below.

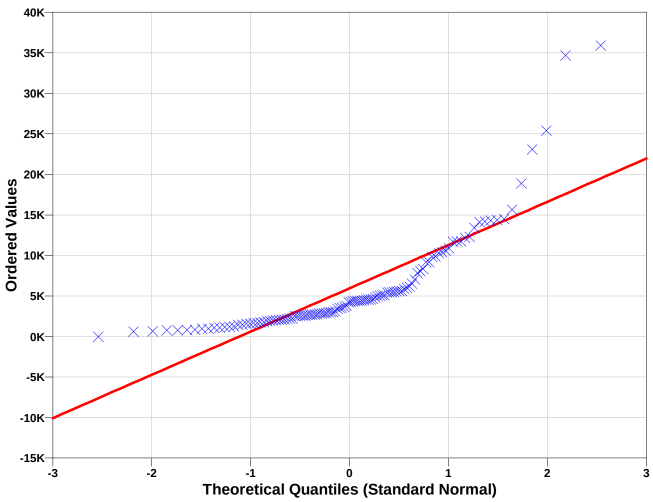
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.2263
Lilliefors 5% Critical Value	0.08372

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	6922

95% Non-Parametric (Chebyshev) UCL	8509

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (8509) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=112 data,
AL is the action level or threshold (6521.2),
SE is the standard error = (standard deviation) / (square root of n).

This *t* was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=111 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-0.9775	1.6587	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
82	65	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

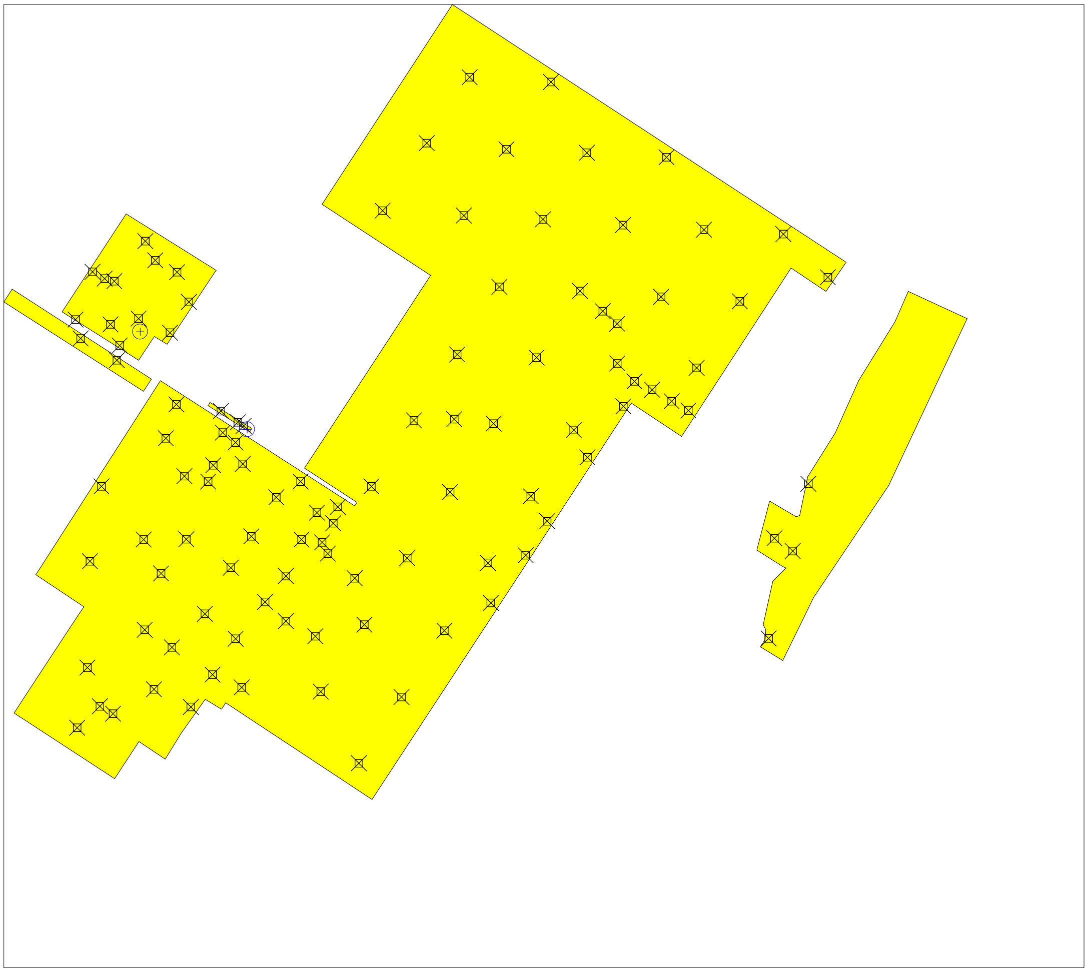
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

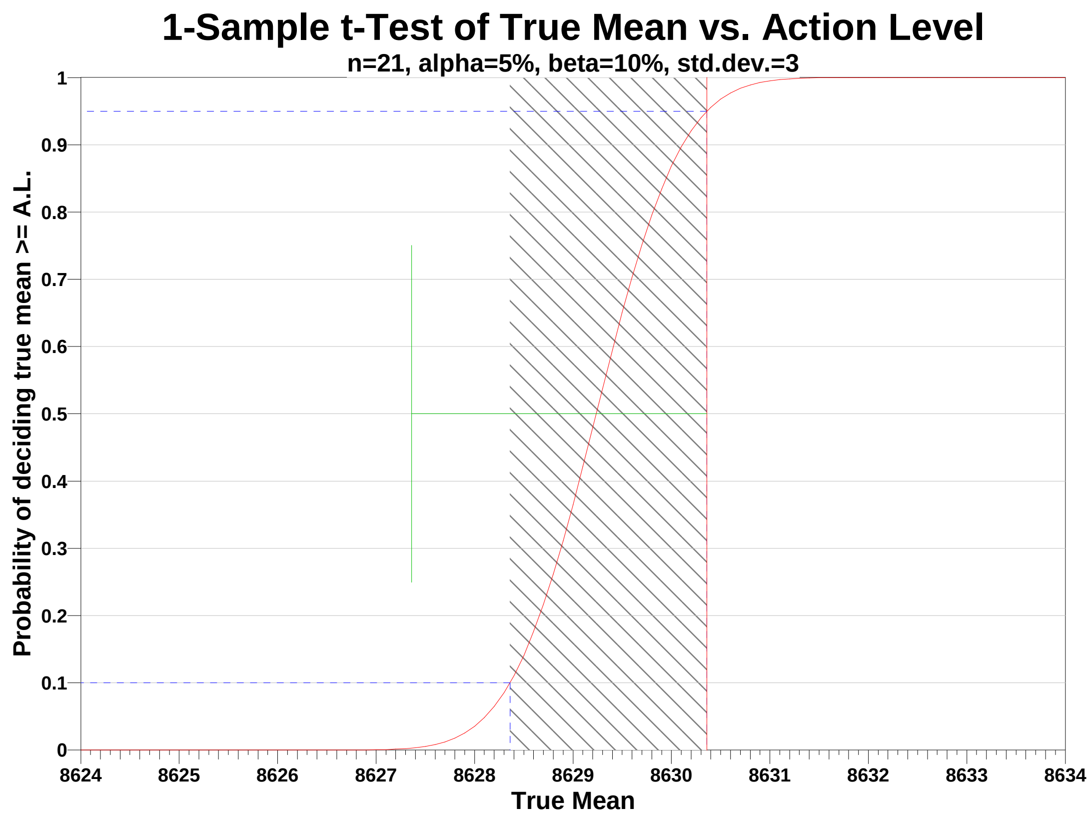
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=8630.36		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	2	2	1	1	1	1
	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1
LBGR=80	β=5	2	2	1	1	1	1
	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1
LBGR=70	β=5	2	2	1	1	1	1

	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	112							
Min	0							
Max	35900							
Range	35900							
Mean	5946.6							
Median	4265							
Variance	3.8705e+007							
StdDev	6221.3							
Std Error	587.86							
Skewness	2.5682							
Interquartile Range	5567.5							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
79.17	808.1	1128	2073	4265	7640	1.389e+004	1.679e+004	3.574e+004

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	4.815	3.414	Yes

The test statistic 4.815 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	35900

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.2158
Lilliefors 5% Critical Value	0.0841

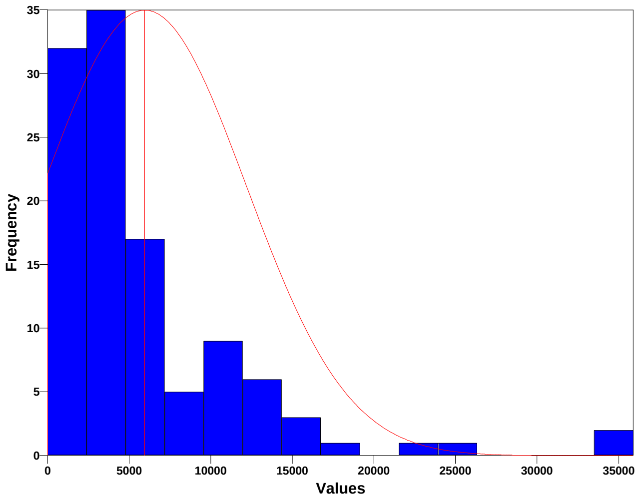
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

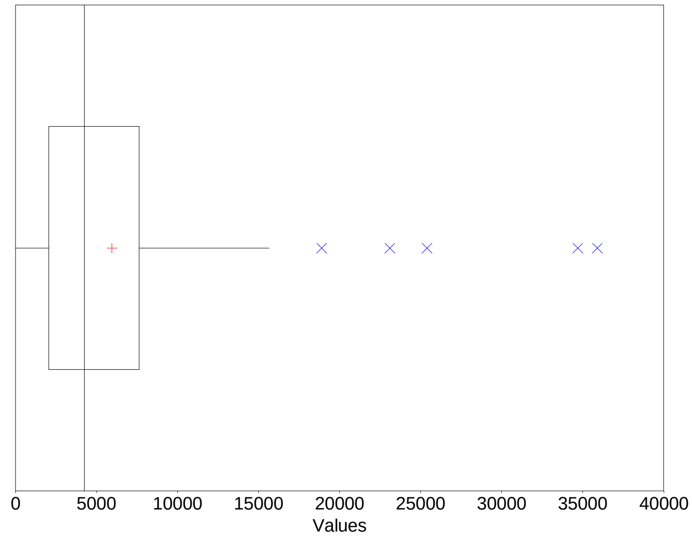
Data Plots
Graphical displays of the data are shown below.

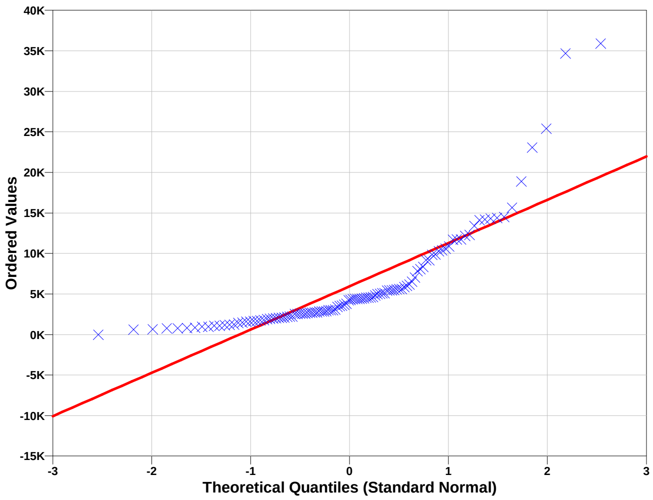
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.2263
Lilliefors 5% Critical Value	0.08372

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	6922

95% Non-Parametric (Chebyshev) UCL	8509

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (8509) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=112 data,
AL is the action level or threshold (8630.36),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=111 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-4.5654	1.6587	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
87	65	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

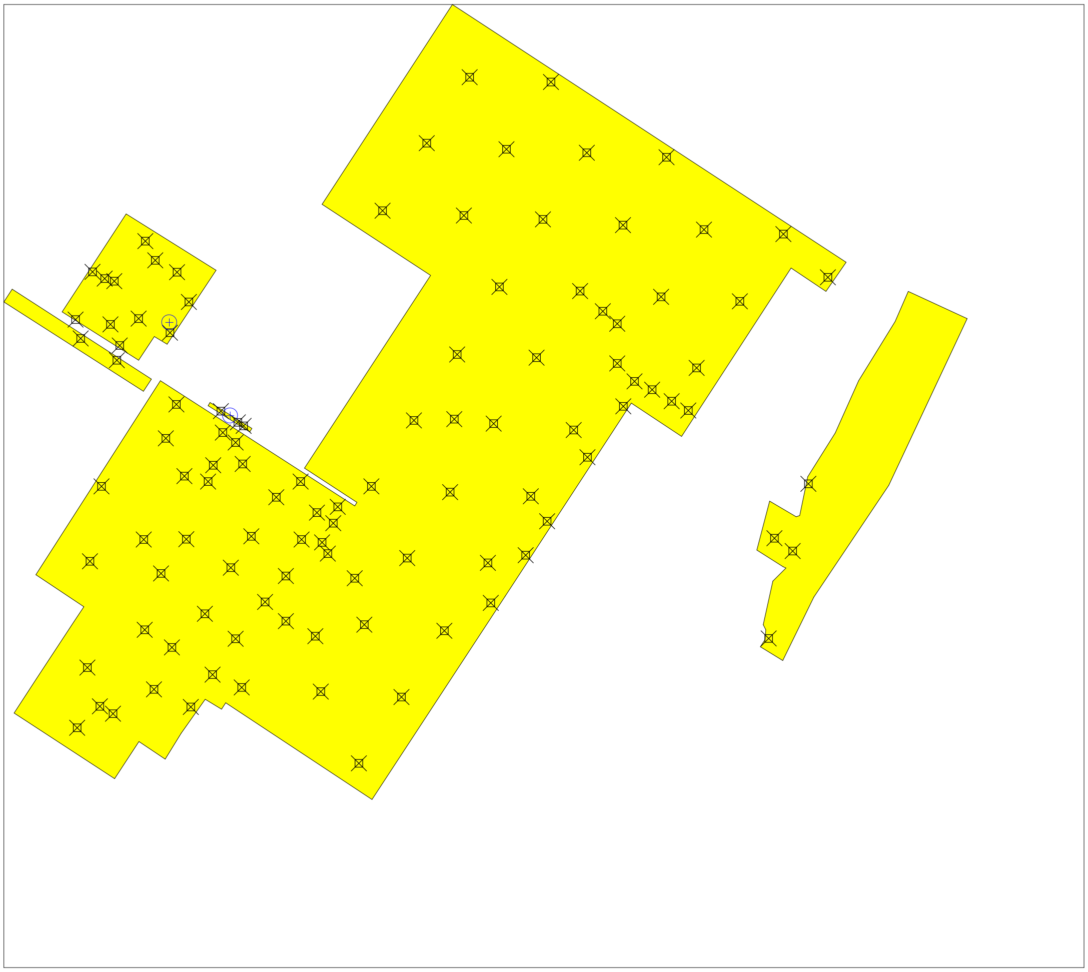
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z^2_{1-\alpha}$$

where

n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

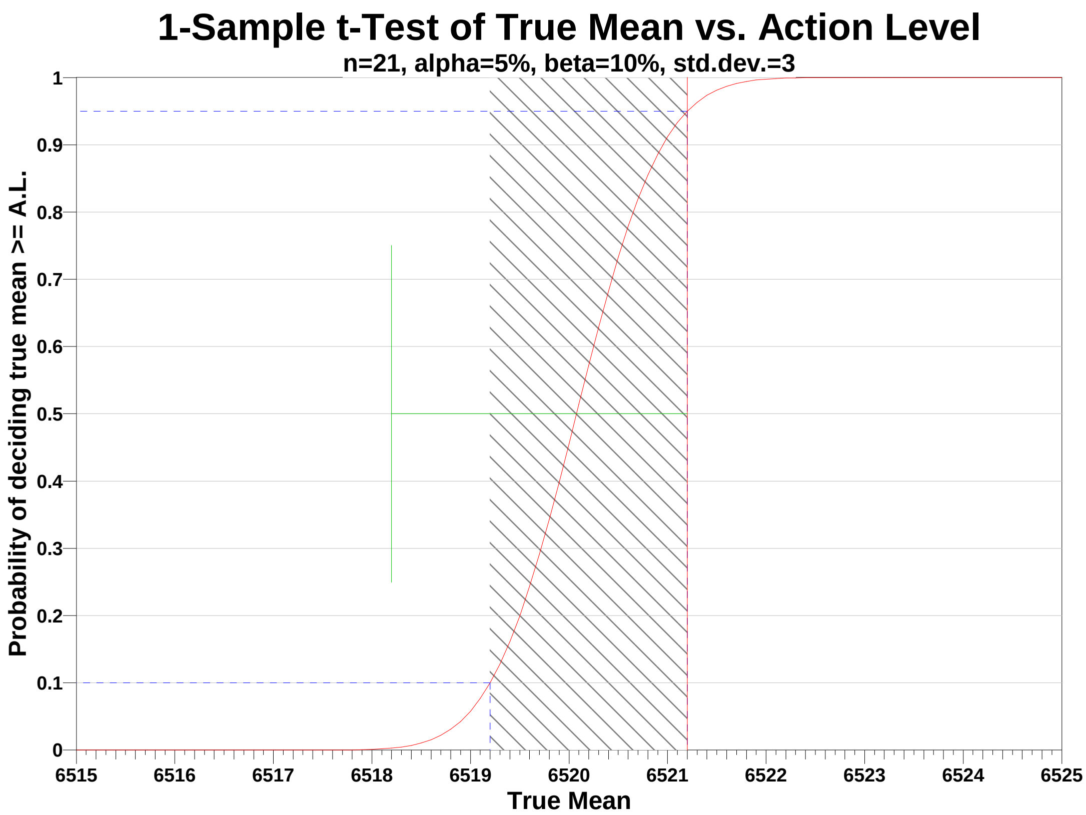
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=6521.2		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	2	2	1	1	1	1
	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1
LBGR=80	β=5	2	2	1	1	1	1
	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1
LBGR=70	β=5	2	2	1	1	1	1

	β=10	2	2	1	1	1	1
	β=15	2	2	1	1	1	1

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	110							
Min	0							
Max	35900							
Range	35900							
Mean	5943.8							
Median	4265							
Variance	3.9239e+007							
StdDev	6264.1							
Std Error	597.26							
Skewness	2.5639							
Interquartile Range	5200							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
66.99	804.7	1113	2060	4265	7260	1.403e+004	1.711e+004	3.577e+004

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	4.764	3.405	Yes

The test statistic 4.764 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	35900

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.2191
Lilliefors 5% Critical Value	0.08526

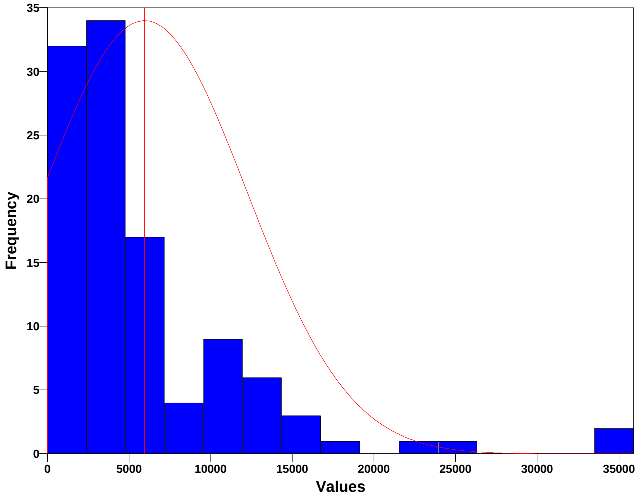
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

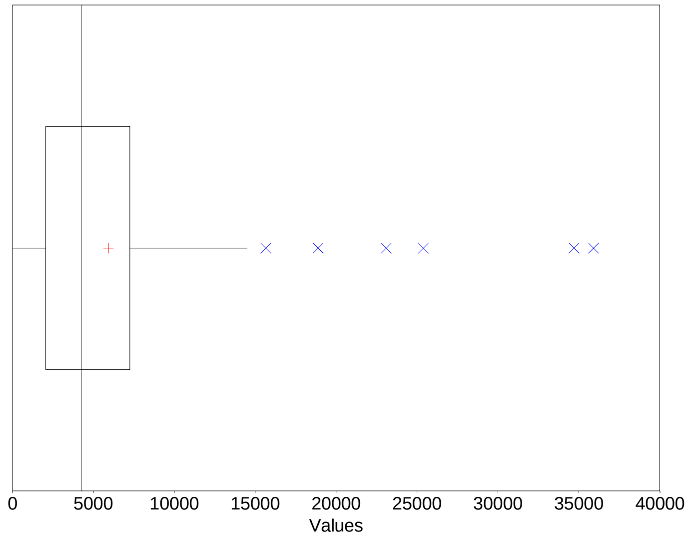
Data Plots
Graphical displays of the data are shown below.

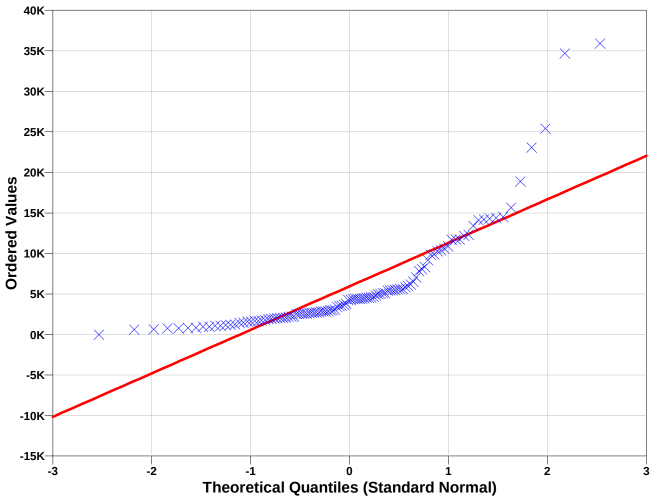
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.2297
Lilliefors 5% Critical Value	0.08448

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	6935

95% Non-Parametric (Chebyshev) UCL	8547

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (8547) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=110 data,
AL is the action level or threshold (6521.2),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=109 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-0.96678	1.659	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
81	64	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

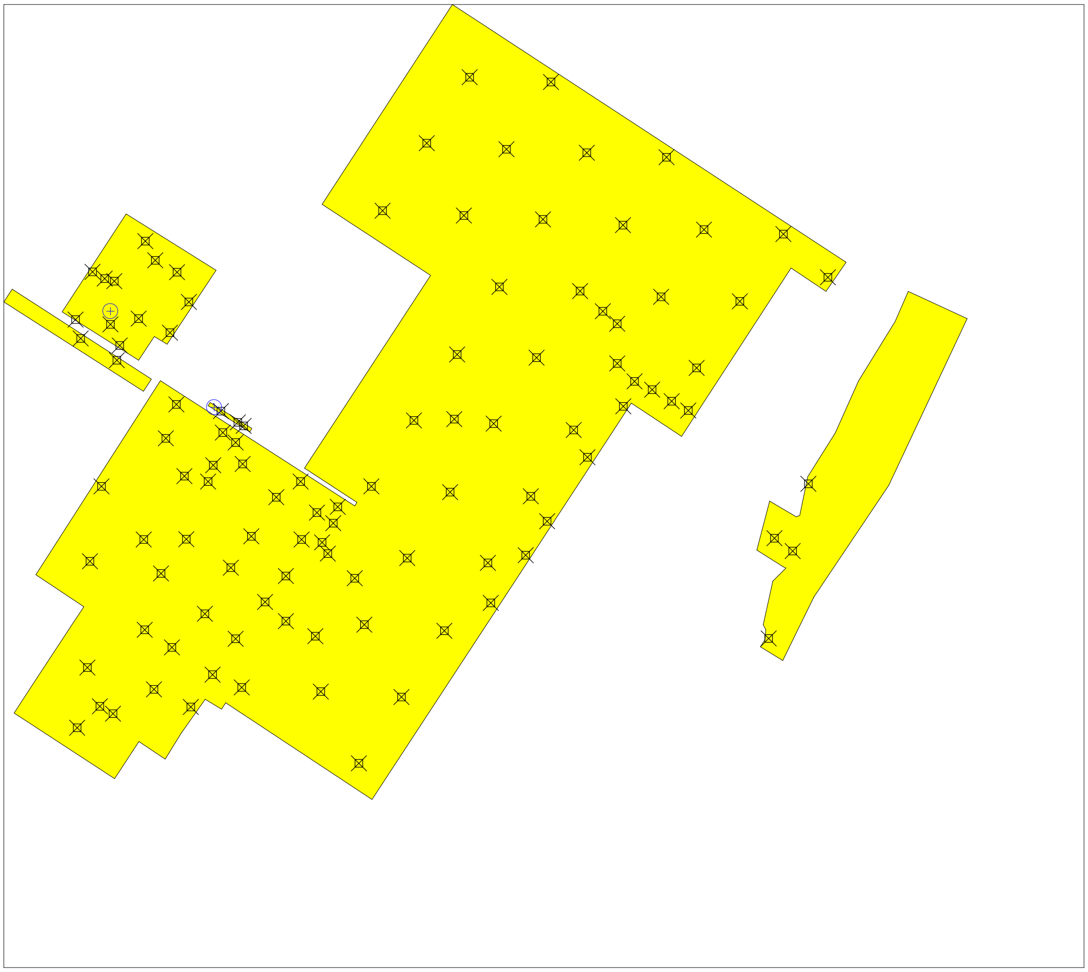
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability (1-β) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is 1-α,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is 1-β.

The values of these inputs that result in the calculated number of sampling locations are:

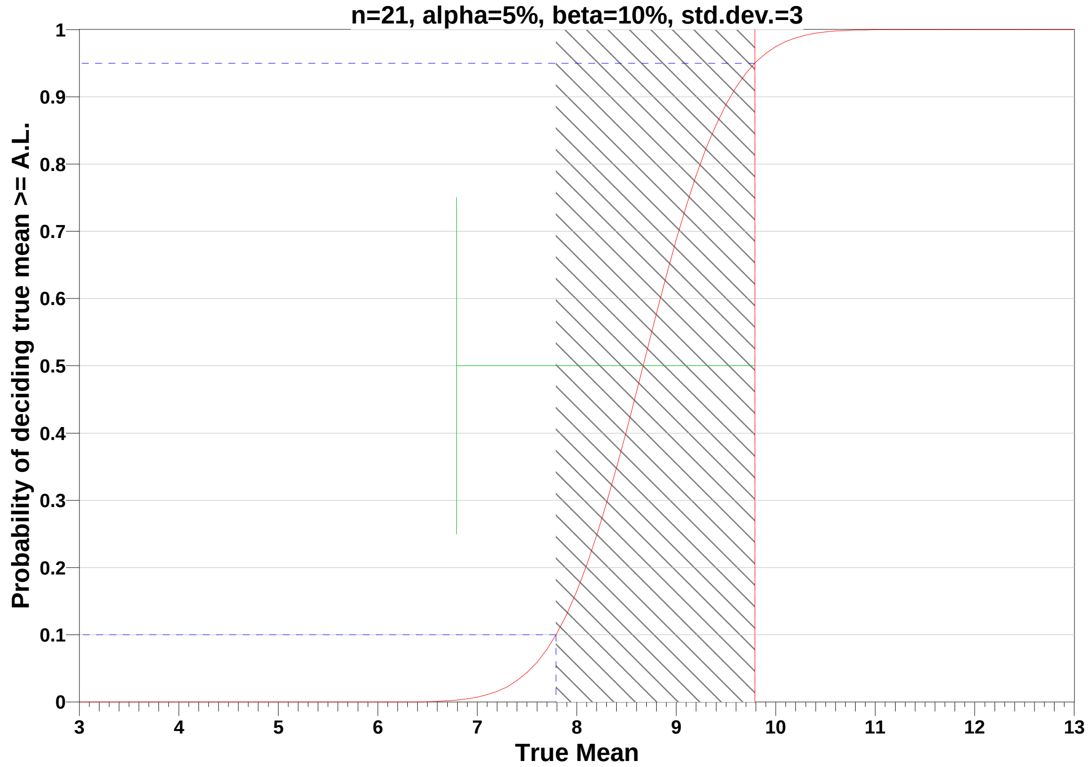
Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

1-Sample t-Test of True Mean vs. Action Level



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=9.79		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	408	103	323	82	271	69
	β=10	324	82	248	63	203	51
	β=15	272	69	203	52	162	41
LBGR=80	**β=5**	103	27	82	21	69	18
	β=10	82	22	63	17	51	14
	β=15	69	19	52	14	41	11
LBGR=70	**β=5**	47	13	37	10	31	9

	β=10	38	11	29	8	23	7
	β=15	32	9	24	7	19	6

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	111							
Min	0							
Max	17.3							
Range	17.3							
Mean	1.8664							
Median	1.3							
Variance	4.868							
StdDev	2.2064							
Std Error	0.20942							
Skewness	3.9354							
Interquartile Range	1.775							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.0216	0.23	0.306	0.625	1.3	2.4	3.28	6.38	16.29

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	6.952	3.405	Yes

The test statistic 6.952 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	17.3

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.1794
Lilliefors 5% Critical Value	0.08526

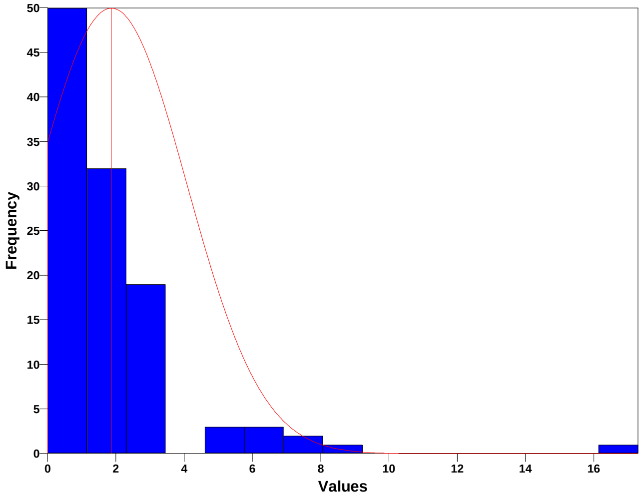
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

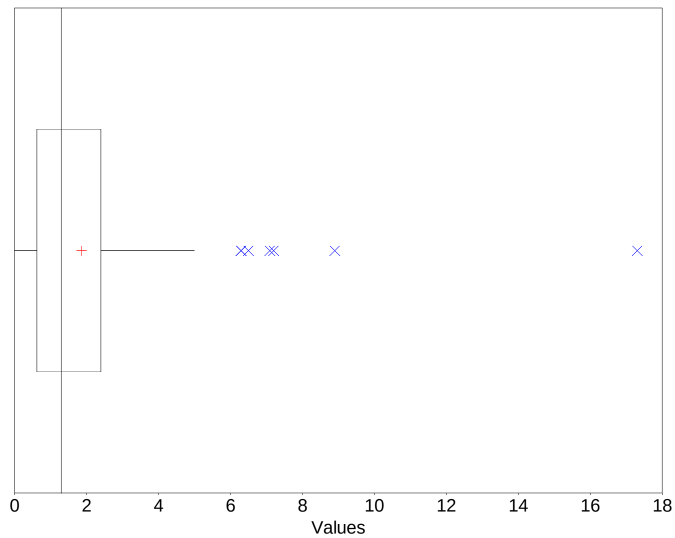
Data Plots
Graphical displays of the data are shown below.

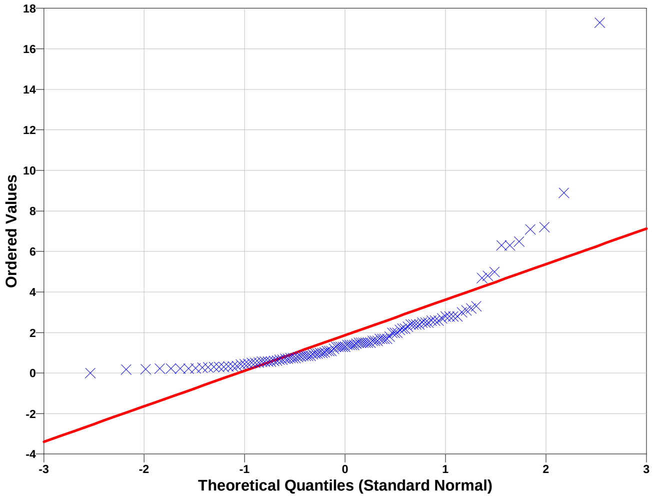
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_n, for which a fraction p of the distribution is less than x_n. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.2133
Lilliefors 5% Critical Value	0.0841

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	2.214

95% Non-Parametric (Chebyshev) UCL	2.779

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (2.779) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=111 data,
AL is the action level or threshold (9.79),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=110 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-37.836	1.6588	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
110	65	Reject

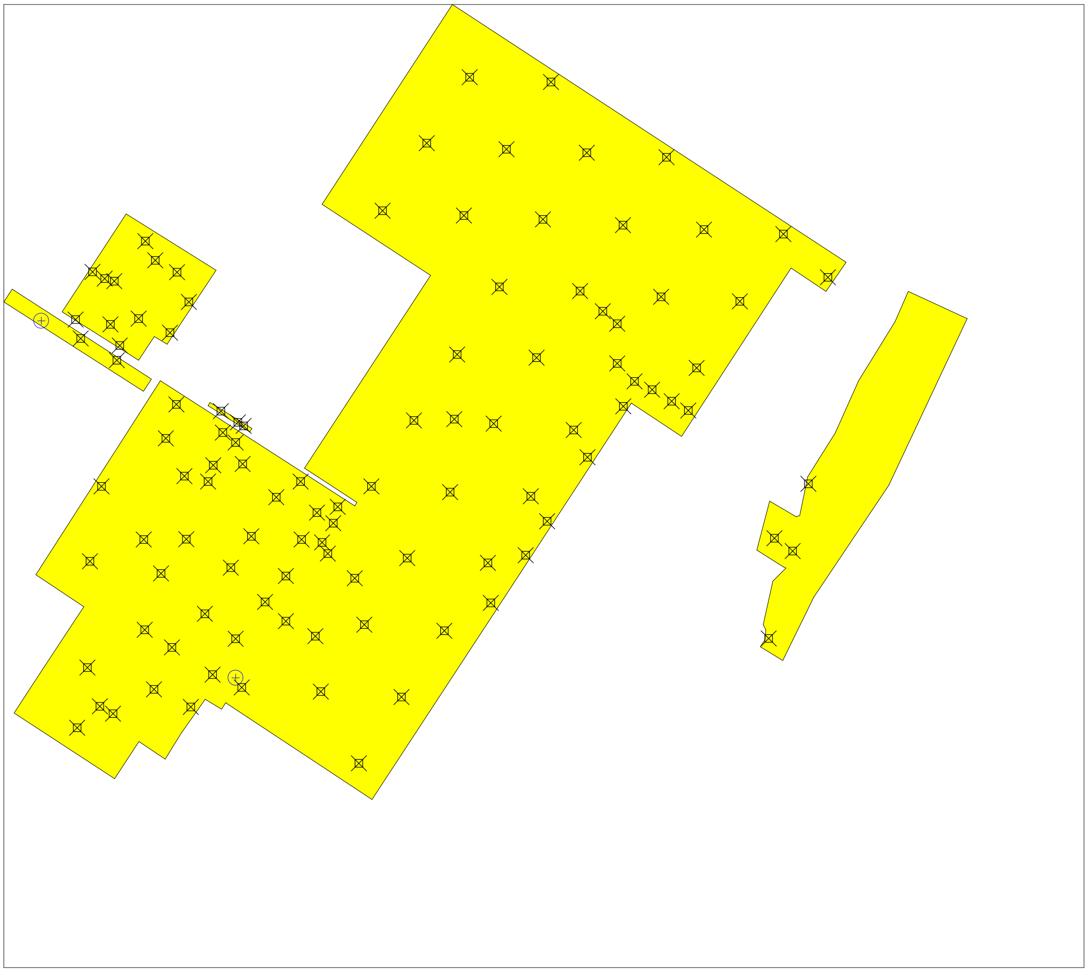
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability (1-β) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

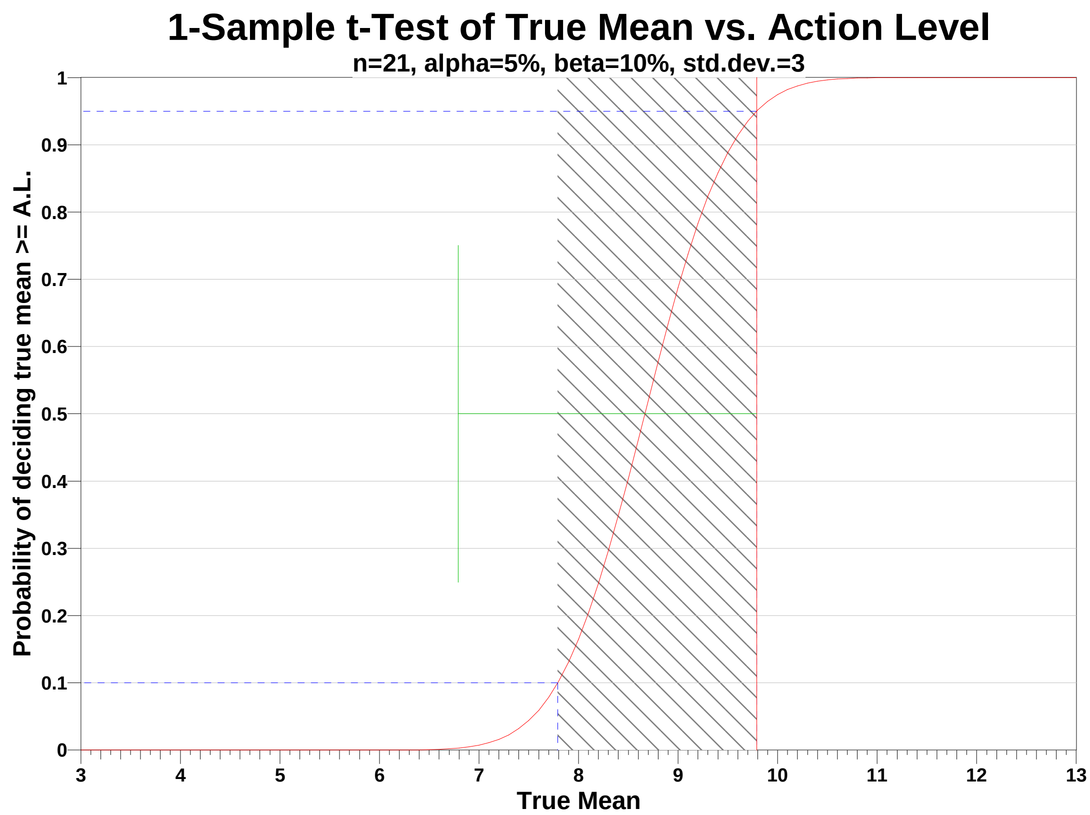
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	**Δ**	**α**	**β**	**$Z_{1-\alpha}$** [a]	**$Z_{1-\beta}$** [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=9.79		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	408	103	323	82	271	69
	β=10	324	82	248	63	203	51
	β=15	272	69	203	52	162	41
LBGR=80	**β=5**	103	27	82	21	69	18
	β=10	82	22	63	17	51	14
	β=15	69	19	52	14	41	11
LBGR=70	**β=5**	47	13	37	10	31	9

	β=10	38	11	29	8	23	7
	β=15	32	9	24	7	19	6

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	112							
Min	0							
Max	17.3							
Range	17.3							
Mean	1.8783							
Median	1.35							
Variance	4.8401							
StdDev	2.2							
Std Error	0.20788							
Skewness	3.9187							
Interquartile Range	1.7662							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.0234	0.23	0.3065	0.6338	1.35	2.4	3.27	6.37	16.21

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	6.995	3.411	Yes

The test statistic 6.995 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	17.3

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.1791
Lilliefors 5% Critical Value	0.08448

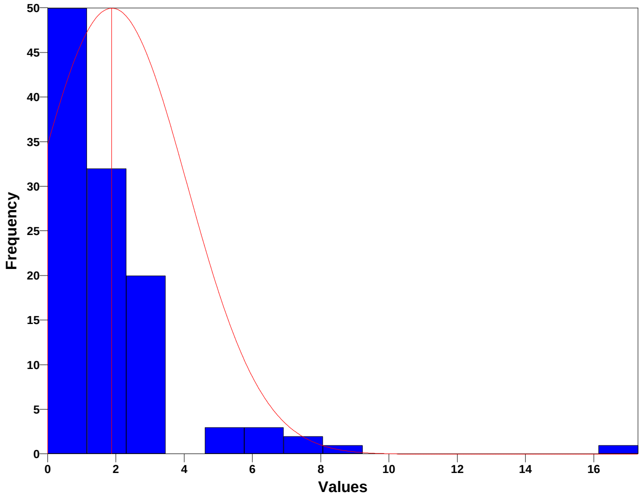
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

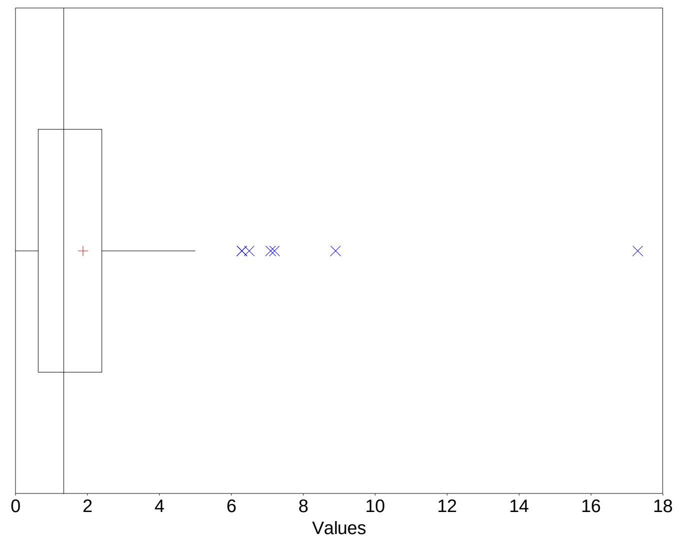
Data Plots
Graphical displays of the data are shown below.

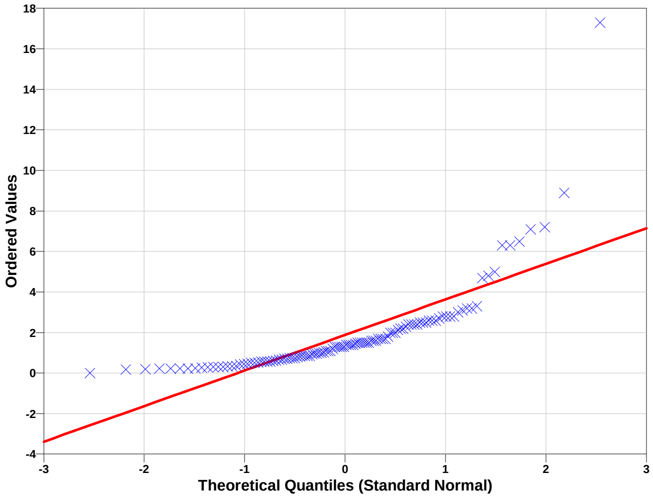
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.2111
Lilliefors 5% Critical Value	0.08372

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	2.223

95% Non-Parametric (Chebyshev) UCL	2.784

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (2.784) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=112 data,
AL is the action level or threshold (9.79),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=111 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-38.058	1.6587	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
111	65	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

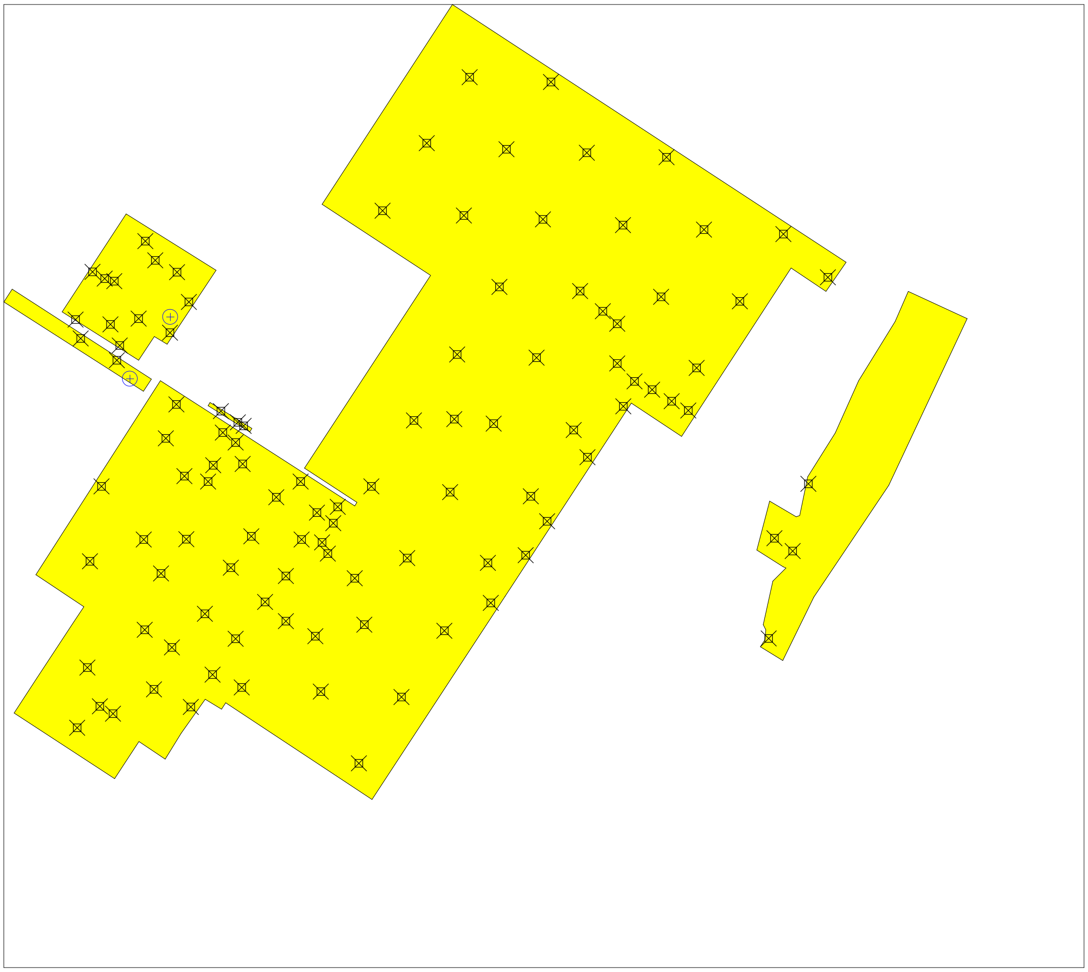
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

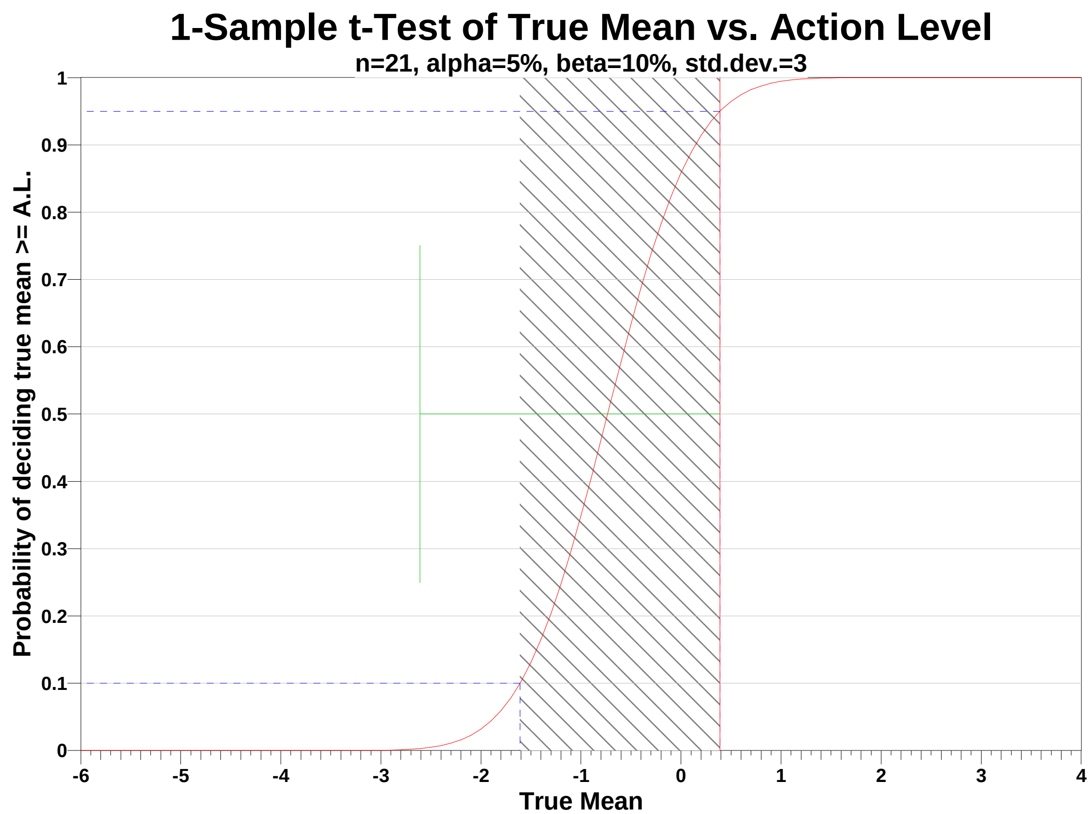
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	**n**	**Parameter**					
		S	**Δ**	**α**	**β**	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.39		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	256148	64038	202696	50675	170162	42541
	β=10	202696	50676	155492	38874	127174	31794
	β=15	170163	42542	127174	31795	101700	25426
LBGR=80	**β=5**	64038	16011	50675	12670	42541	10636
	β=10	50676	12670	38874	9720	31794	7949
	β=15	42542	10637	31795	7950	25426	6357
LBGR=70	**β=5**	28463	7117	22523	5632	18908	4728

	β=10	22523	5632	17278	4321	14131	3534
	β=15	18909	4729	14132	3534	11301	2826

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	110							
Min	0							
Max	17.3							
Range	17.3							
Mean	1.8782							
Median	1.35							
Variance	4.8971							
StdDev	2.2129							
Std Error	0.211							
Skewness	3.9224							
Interquartile Range	1.7487							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.0198	0.23	0.3055	0.6513	1.35	2.4	3.29	6.39	16.38

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	6.969	3.408	Yes

The test statistic 6.969 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	17.3

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.1786
Lilliefors 5% Critical Value	0.08486

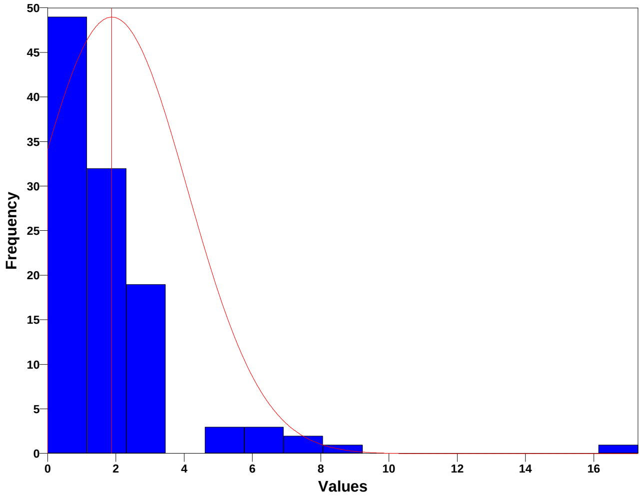
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

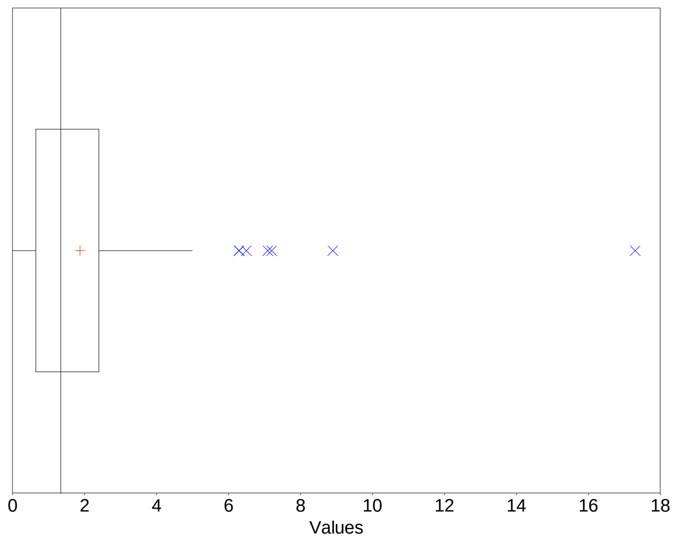
Data Plots
Graphical displays of the data are shown below.

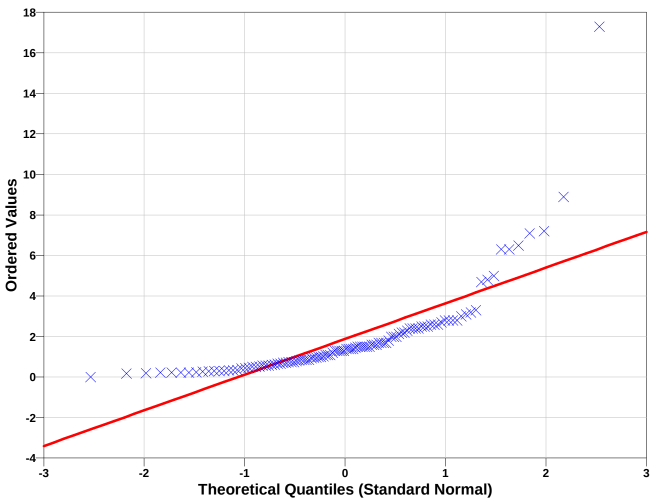
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.2123
Lilliefors 5% Critical Value	0.08448

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	2.228

95% Non-Parametric (Chebyshev) UCL	2.798

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (2.798) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=110 data,
AL is the action level or threshold (0.39),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=109 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
7.0532	1.659	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
16	64	Cannot Reject
Note: There may not be enough data to reject the null hypothesis (and conclude site is clean) with 95% confidence using the MARSSIM sign test.		

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

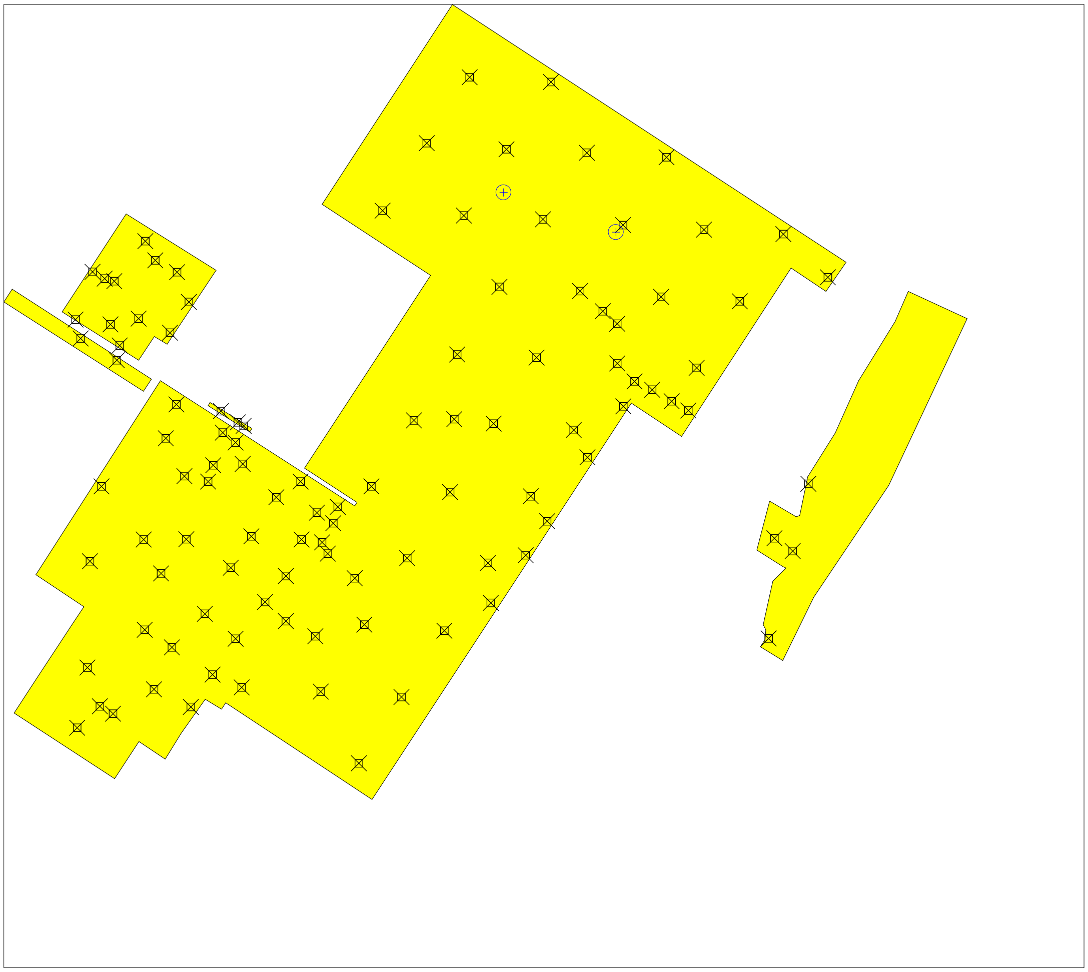
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability (1-β) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5 Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is 1-α,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is 1-β.

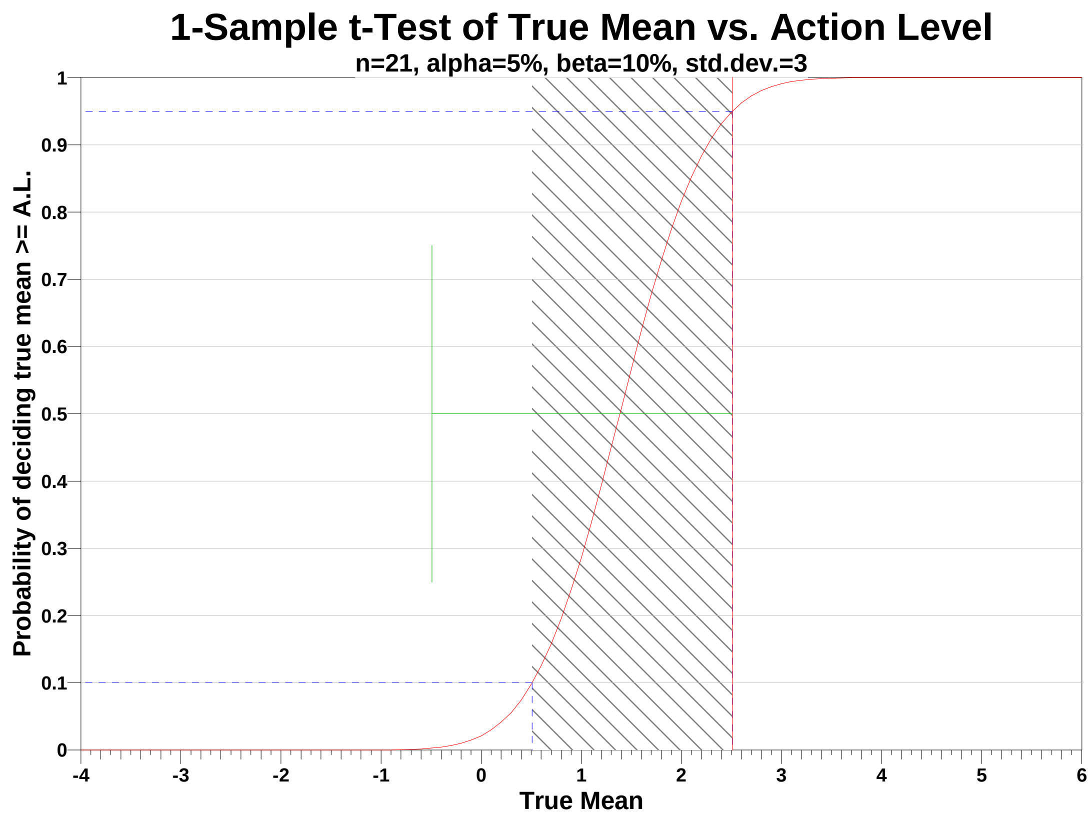
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	**Δ**	**α**	**β**	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=2.50958		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	6188	1548	4897	1225	4111	1028
	β=10	4897	1226	3757	940	3072	769
	β=15	4111	1029	3073	769	2457	615
LBGR=80	**β=5**	1548	388	1225	307	1028	258
	β=10	1226	308	940	236	769	193
	β=15	1029	259	769	193	615	155
LBGR=70	**β=5**	689	174	545	137	458	115

	β=10	546	138	419	106	342	86
	β=15	458	116	343	87	274	69

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n				114				
Min				0				
Max				17.3				
Range				17.3				
Mean				1.9277				
Median				1.4				
Variance				5.0154				
StdDev				2.2395				
Std Error				0.20975				
Skewness				3.6923				
Interquartile Range				1.7737				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.2275	0.2975	0.6513	1.4	2.425	4	6.65	16.04

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	7.023	3.416	Yes

The test statistic 7.023 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	17.3

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.1715
Lilliefors 5% Critical Value	0.08372

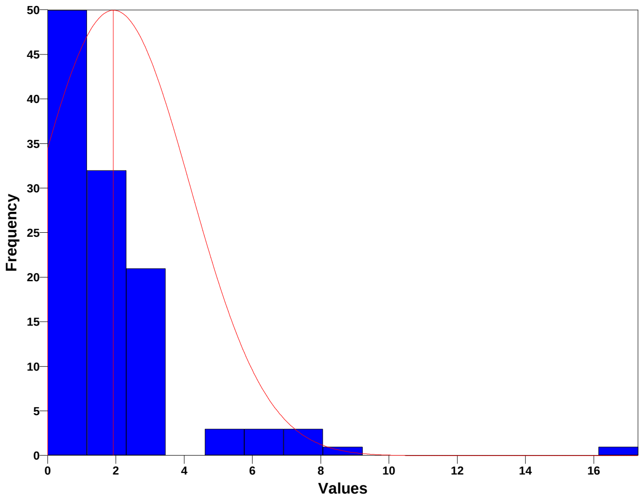
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

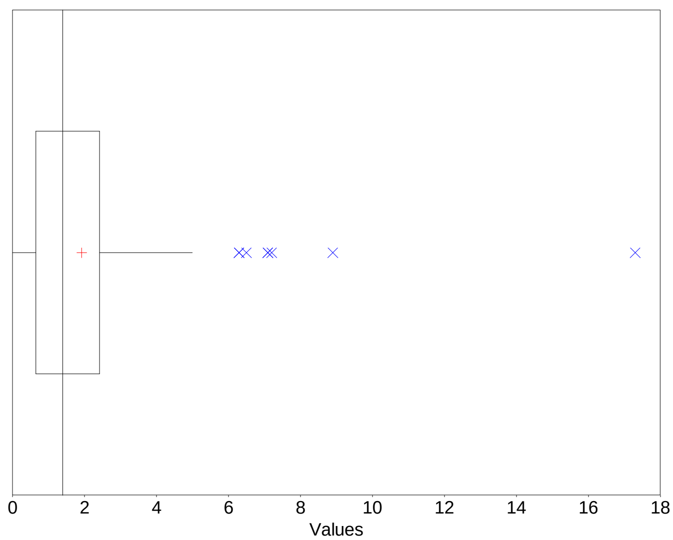
Data Plots
Graphical displays of the data are shown below.

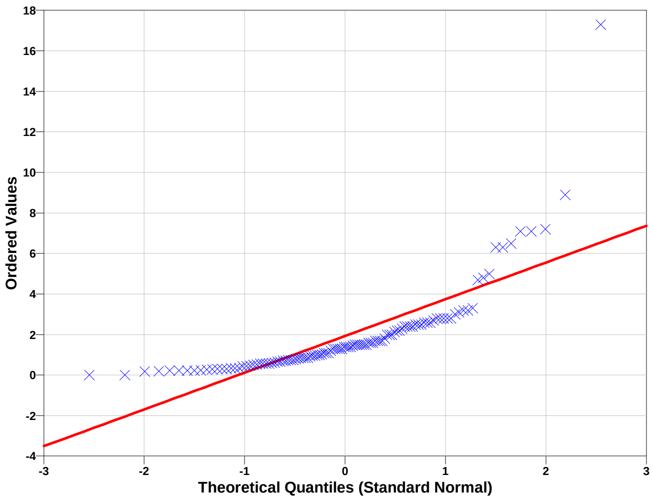
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.2081
Lilliefors 5% Critical Value	0.08298

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	2.276

95% Non-Parametric (Chebyshev) UCL	2.842

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (2.842) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value t was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=114 data,
AL is the action level or threshold (2.50958),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=113 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-2.774	1.6585	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
89	66	Reject

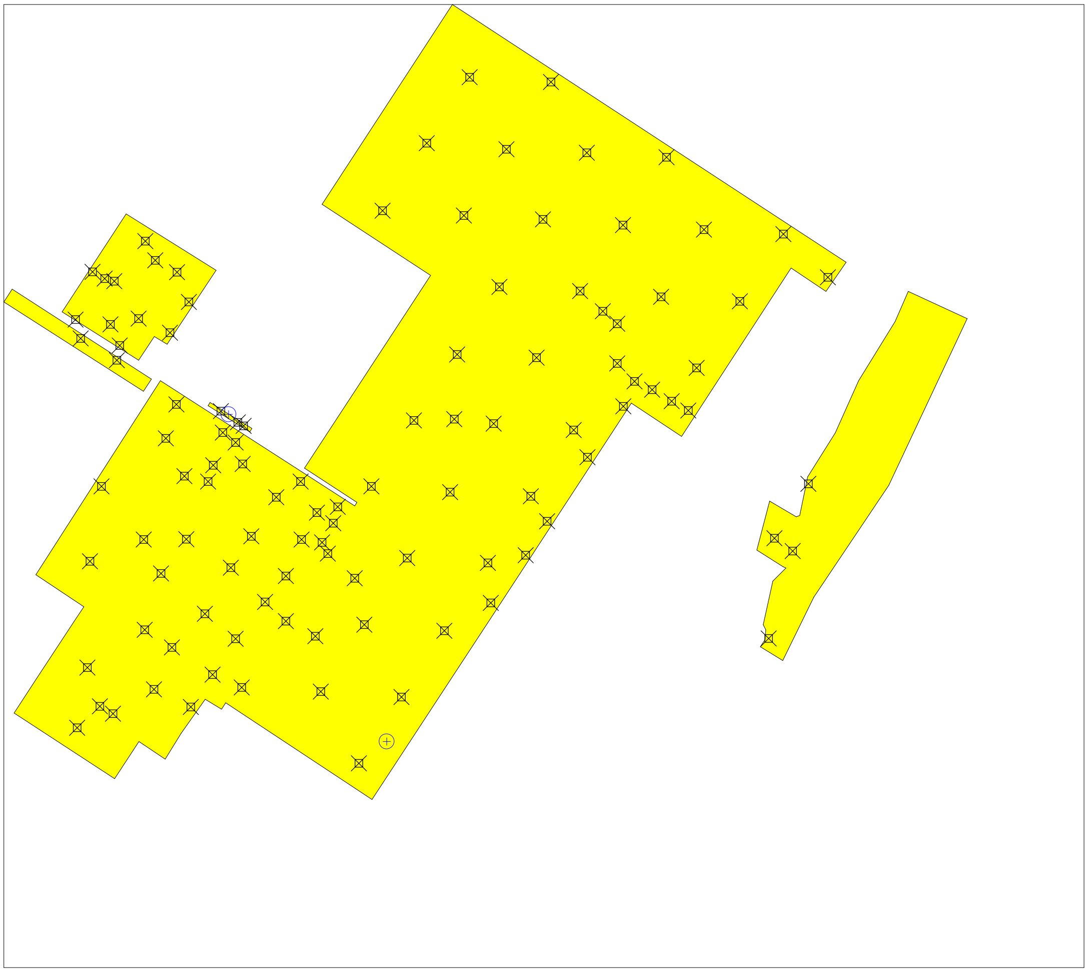
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

The values of these inputs that result in the calculated number of sampling locations are:

Analyte	**n**	**Parameter**					
		S	**Δ**	**α**	**β**	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.

Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=24.2		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	68	18	54	14	45	12
	β=10	54	15	42	11	34	9
	β=15	46	13	34	10	27	8
LBGR=80	β=5	18	6	14	5	12	4
	β=10	15	5	11	4	9	3
	β=15	13	5	10	3	8	3
LBGR=70	β=5	9	4	7	3	6	2

	β=10	8	3	6	2	5	2
	β=15	7	3	5	2	4	2

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	111							
Min	0							
Max	17.3							
Range	17.3							
Mean	1.8771							
Median	1.3							
Variance	4.8978							
StdDev	2.2131							
Std Error	0.21006							
Skewness	3.8834							
Interquartile Range	1.775							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.226	0.293	0.625	1.3	2.4	3.28	6.38	16.29

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	6.928	3.405	Yes

The test statistic 6.928 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	17.3

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.1749
Lilliefors 5% Critical Value	0.08526

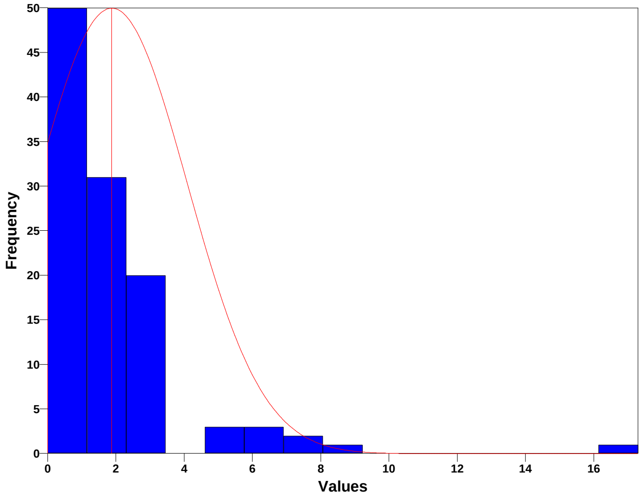
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

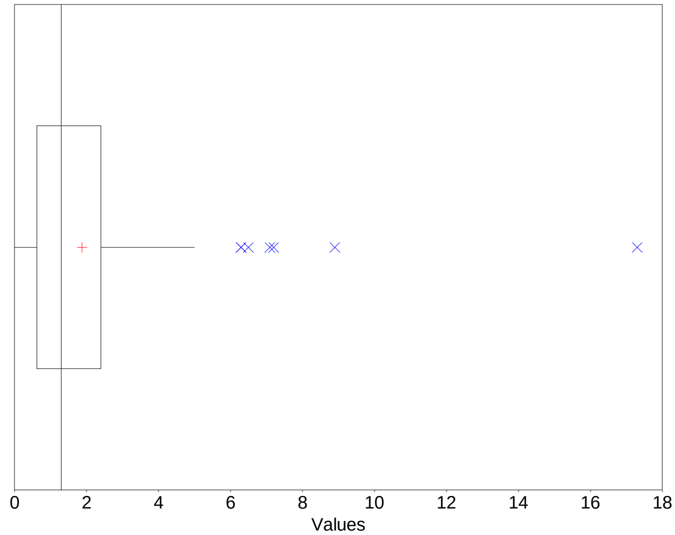
Data Plots
Graphical displays of the data are shown below.

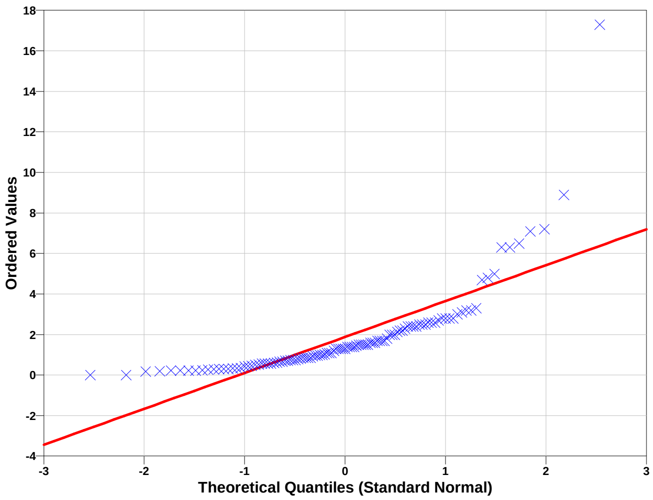
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.2036
Lilliefors 5% Critical Value	0.0841

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	2.226

95% Non-Parametric (Chebyshev) UCL	2.793

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (2.793) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=111 data,
AL is the action level or threshold (24.2),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=110 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-106.27	1.6588	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
111	65	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

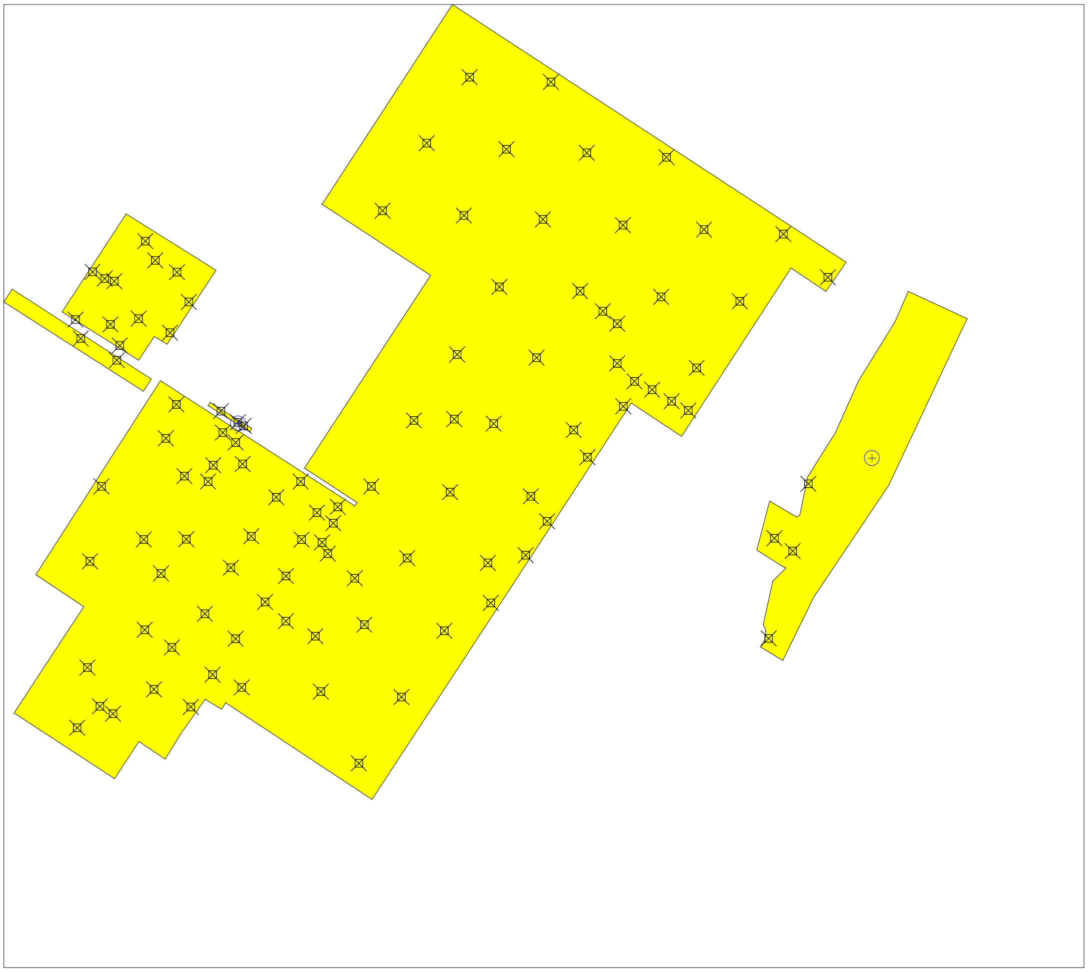
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability (1-β) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

The values of these inputs that result in the calculated number of sampling locations are:

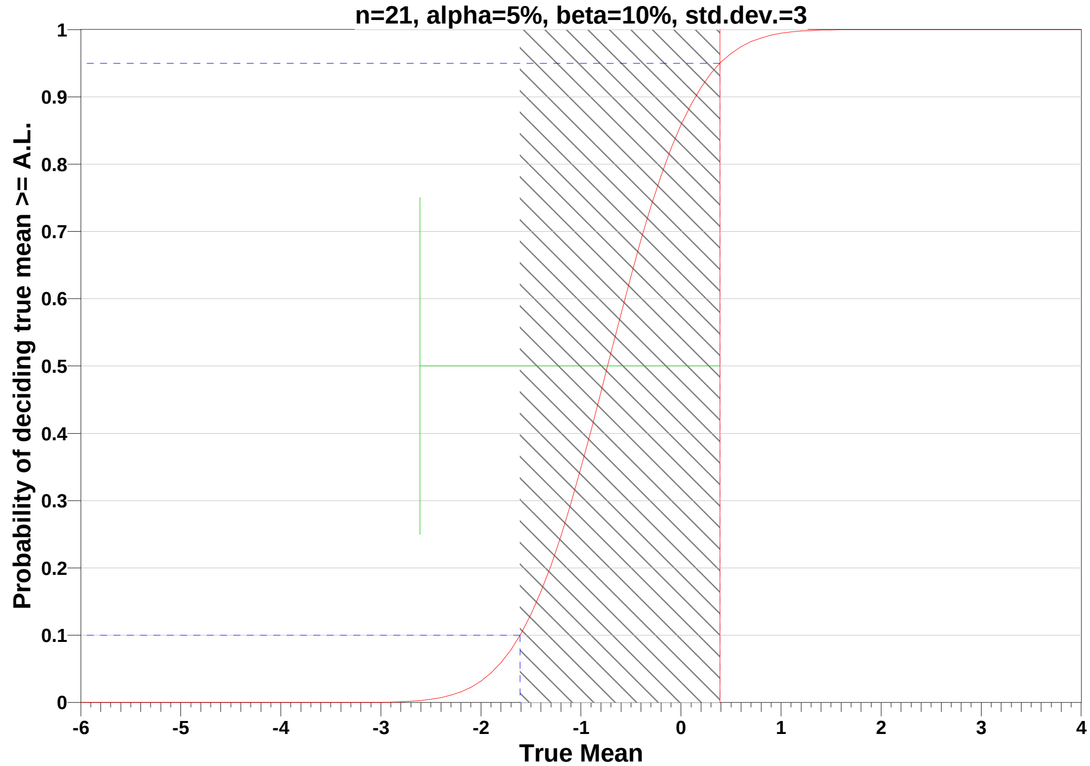
Analyte	**n**	**Parameter**					
		S	**Δ**	**α**	**β**	***Z*$_{1-\alpha}$** [a]	***Z*$_{1-\beta}$** [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

1-Sample t-Test of True Mean vs. Action Level



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.39		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	256148	64038	202696	50675	170162	42541
	β=10	202696	50676	155492	38874	127174	31794
	β=15	170163	42542	127174	31795	101700	25426
LBGR=80	**β=5**	64038	16011	50675	12670	42541	10636
	β=10	50676	12670	38874	9720	31794	7949
	β=15	42542	10637	31795	7950	25426	6357
LBGR=70	**β=5**	28463	7117	22523	5632	18908	4728

	β=10	22523	5632	17278	4321	14131	3534
	β=15	18909	4729	14132	3534	11301	2826

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	109							
Min	0.09							
Max	17.3							
Range	17.21							
Mean	1.8938							
Median	1.4							
Variance	4.9227							
StdDev	2.2187							
Std Error	0.21251							
Skewness	3.9092							
Interquartile Range	1.735							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.0905	0.17	0.31	0.665	1.4	2.4	3.3	6.4	16.46

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	6.885	3.4	Yes

The test statistic 6.885 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	17.3

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.1759
Lilliefors 5% Critical Value	0.08606

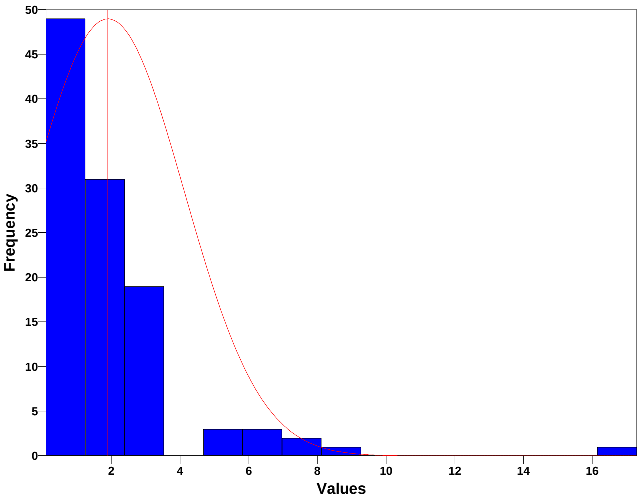
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

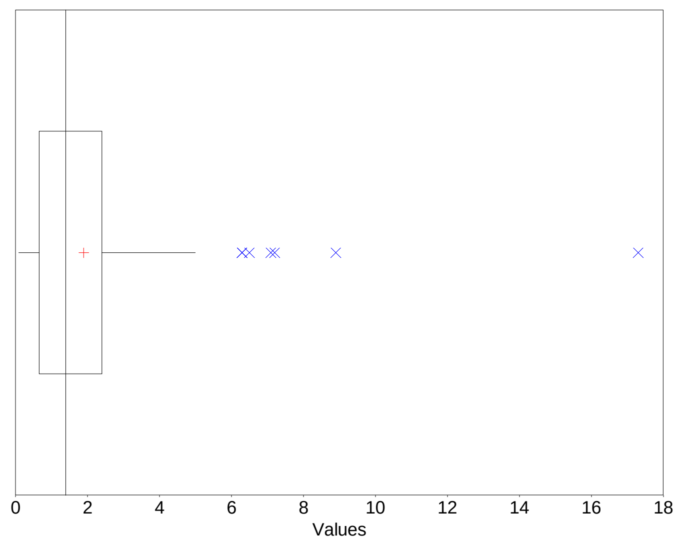
Data Plots
Graphical displays of the data are shown below.

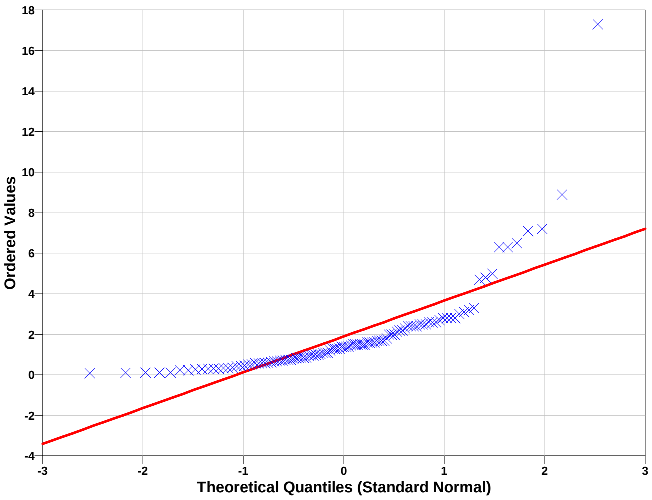
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.213
Lilliefors 5% Critical Value	0.08486

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	2.246

95% Non-Parametric (Chebyshev) UCL	2.82

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (2.82) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=109 data,
AL is the action level or threshold (0.39),
SE is the standard error = (standard deviation) / (square root of n).

This *t* was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=108 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
7.0764	1.6591	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
15	64	Cannot Reject
Note: There may not be enough data to reject the null hypothesis (and conclude site is clean) with 95% confidence using the MARSSIM sign test.		

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

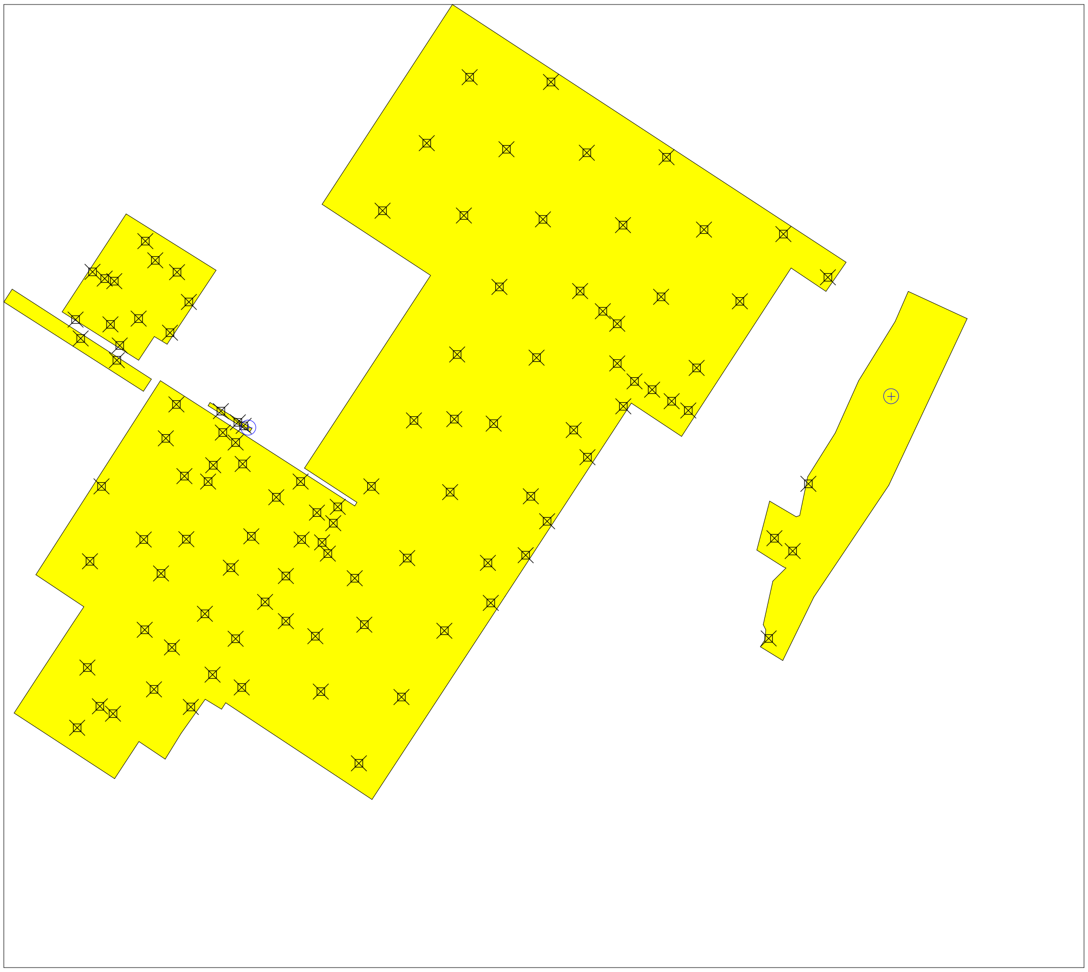
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

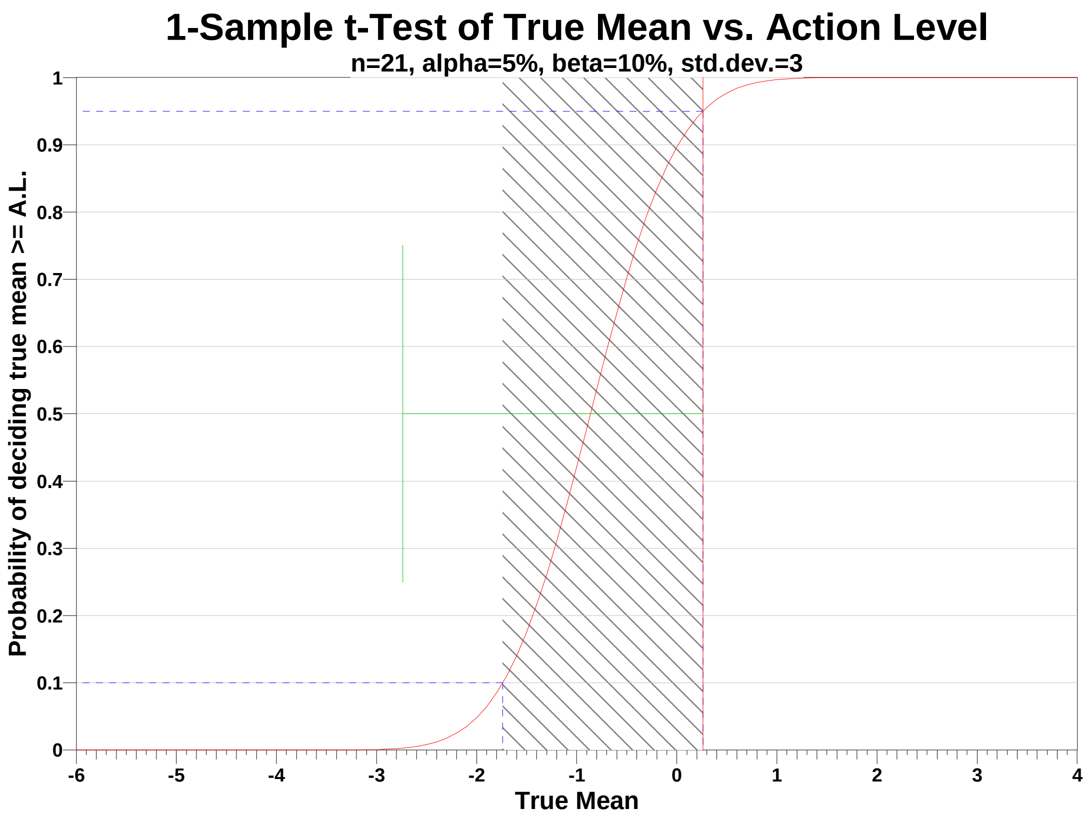
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.261		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	571923	142982	452576	113145	379935	94985
	β=10	452577	113146	347180	86796	283952	70989
	β=15	379936	94985	283952	70989	227073	56769
LBGR=80	β=5	142982	35747	113145	28287	94985	23747
	β=10	113146	28288	86796	21700	70989	17748
	β=15	94985	23748	70989	17748	56769	14193
LBGR=70	β=5	63549	15889	50287	12573	42216	10555

	β=10	50288	12573	38577	9645	31551	7889
	β=15	42217	10556	31551	7889	25231	6309

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is \$11,500.00, which averages out to a per sample cost of \$547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		\$100.00	\$2,100.00
Analytical costs	\$400.00	\$400.00	\$8,400.00
Sum of Field & Analytical costs		**\$500.00**	**\$10,500.00**
Fixed planning and validation costs			\$1,000.00
Total cost			**\$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n				113				
Min				0				
Max				3.97				
Range				3.97				
Mean				0.18366				
Median				0.079				
Variance				0.18895				
StdDev				0.43469				
Std Error				0.040892				
Skewness				6.8685				
Interquartile Range				0.024				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.00868	0.068	0.07	0.074	0.079	0.098	0.374	0.726	3.704

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	8.711	3.416	Yes

The test statistic 8.711 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	3.97

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.3975
Lilliefors 5% Critical Value	0.08372

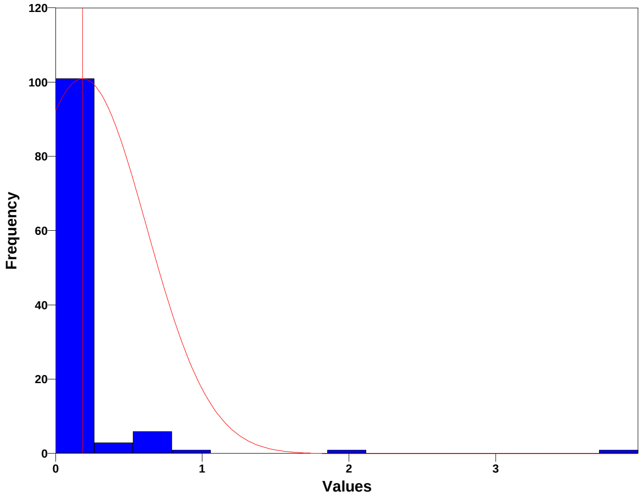
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

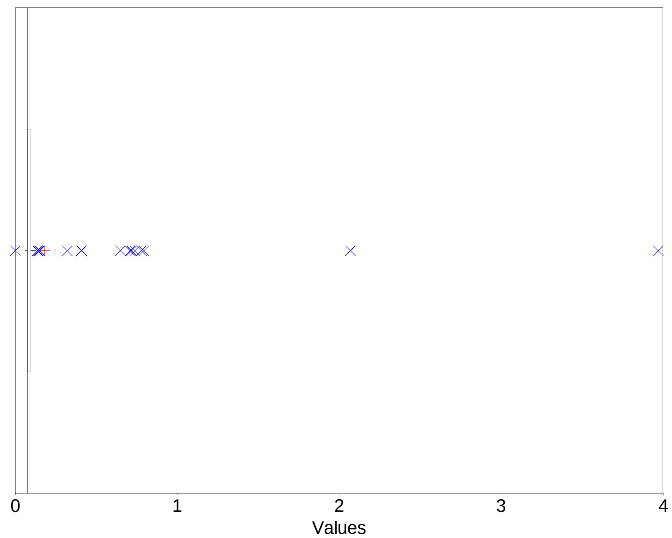
Data Plots
Graphical displays of the data are shown below.

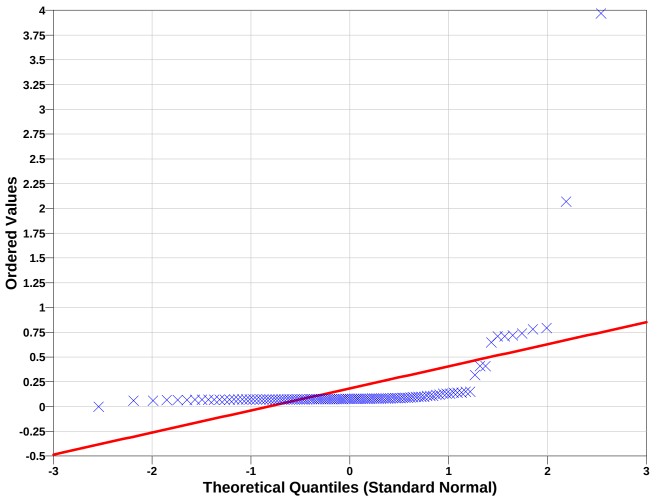
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.4224
Lilliefors 5% Critical Value	0.08335

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.2515

95% Non-Parametric (Chebyshev) UCL	0.3619

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.3619) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=113 data,
AL is the action level or threshold (0.261),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=112 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-1.8912	1.6586	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
101	66	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

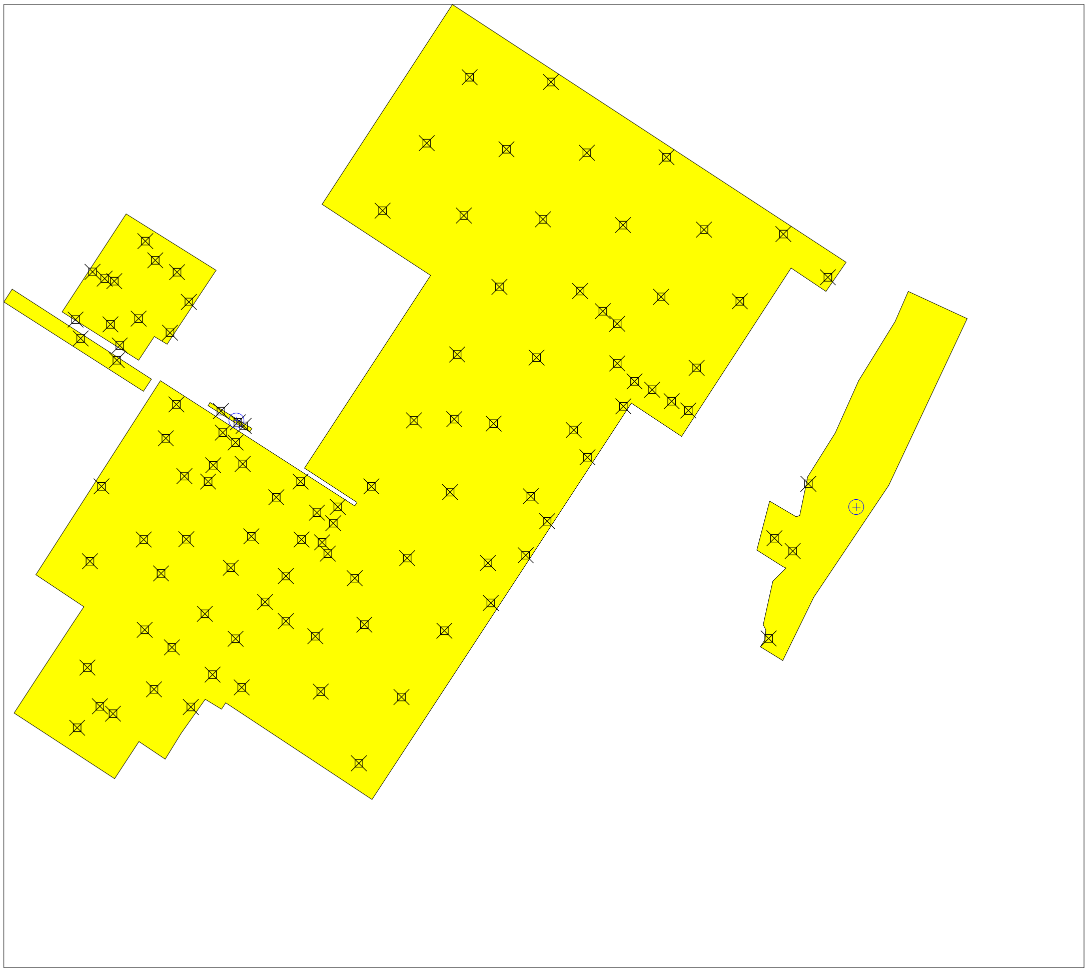
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

- n is the number of samples,
- S is the estimated standard deviation of the measured values including analytical error,
- Δ is the width of the gray region,
- α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
- β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
- $Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
- $Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

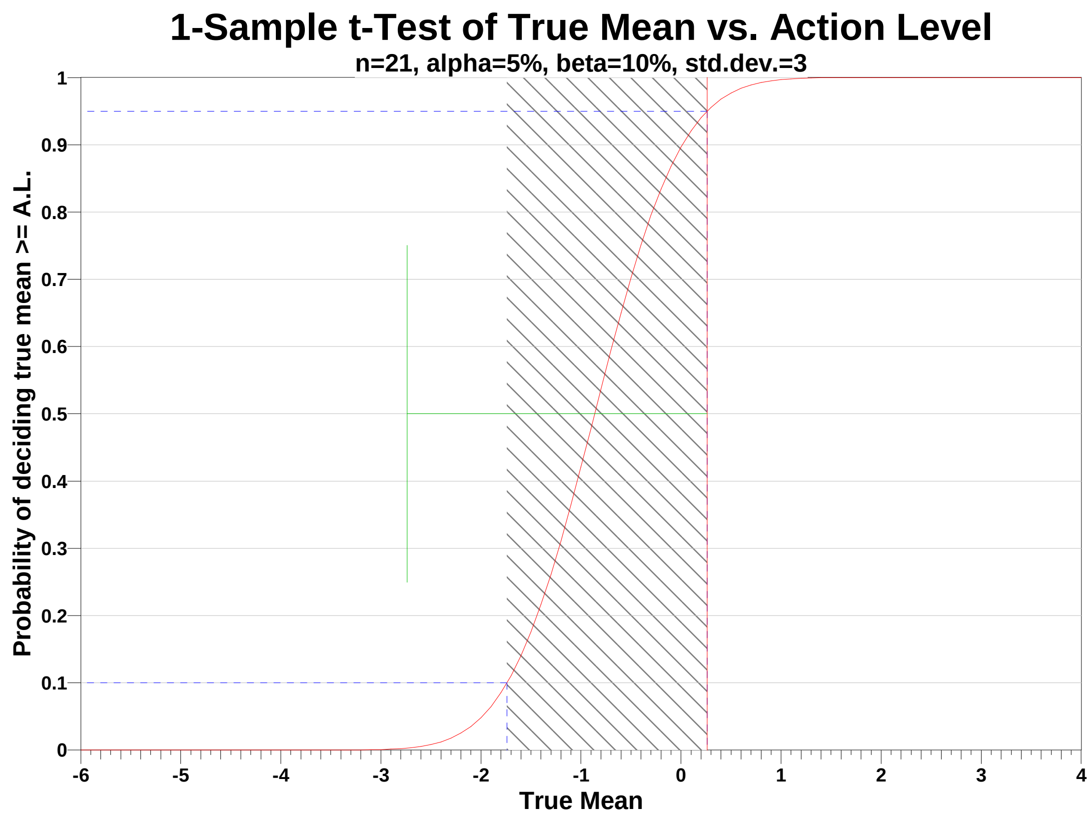
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	**n**	**Parameter**					
		S	**Δ**	**α**	**β**	**$Z_{1-\alpha}$** [a]	**$Z_{1-\beta}$** [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.261		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	571923	142982	452576	113145	379935	94985
	β=10	452577	113146	347180	86796	283952	70989
	β=15	379936	94985	283952	70989	227073	56769
LBGR=80	β=5	142982	35747	113145	28287	94985	23747
	β=10	113146	28288	86796	21700	70989	17748
	β=15	94985	23748	70989	17748	56769	14193
LBGR=70	β=5	63549	15889	50287	12573	42216	10555

	β=10	50288	12573	38577	9645	31551	7889
	β=15	42217	10556	31551	7889	25231	6309

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	112							
Min	0							
Max	3.97							
Range	3.97							
Mean	0.18468							
Median	0.07925							
Variance	0.19054							
StdDev	0.43651							
Std Error	0.041246							
Skewness	6.8383							
Interquartile Range	0.024							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.00806	0.068	0.0703	0.074	0.07925	0.098	0.383	0.727	3.723

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	8.672	3.414	Yes

The test statistic 8.672 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	3.97

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.3978
Lilliefors 5% Critical Value	0.0841

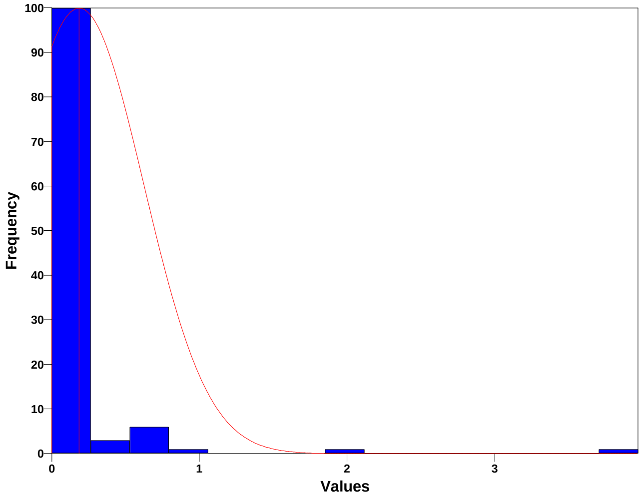
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

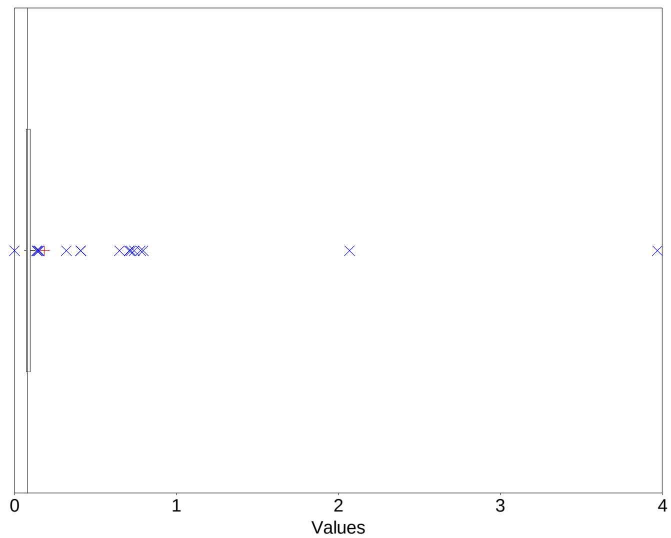
Data Plots
Graphical displays of the data are shown below.

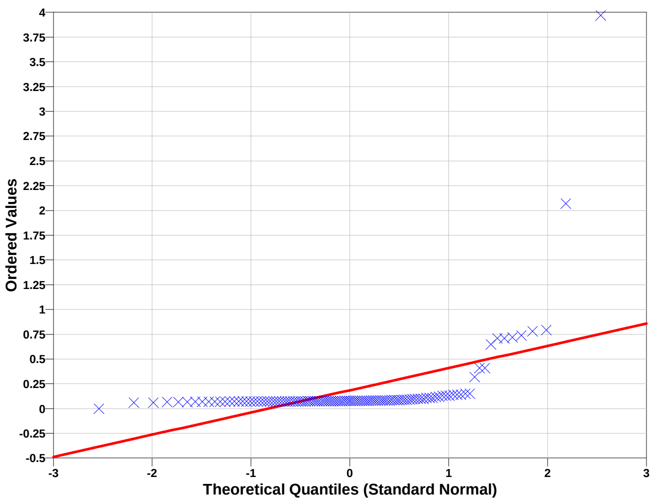
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.4222
Lilliefors 5% Critical Value	0.08372

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.2531

95% Non-Parametric (Chebyshev) UCL	0.3645

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.3645) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=112 data,
AL is the action level or threshold (0.261),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=111 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-1.8504	1.6587	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
100	65	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

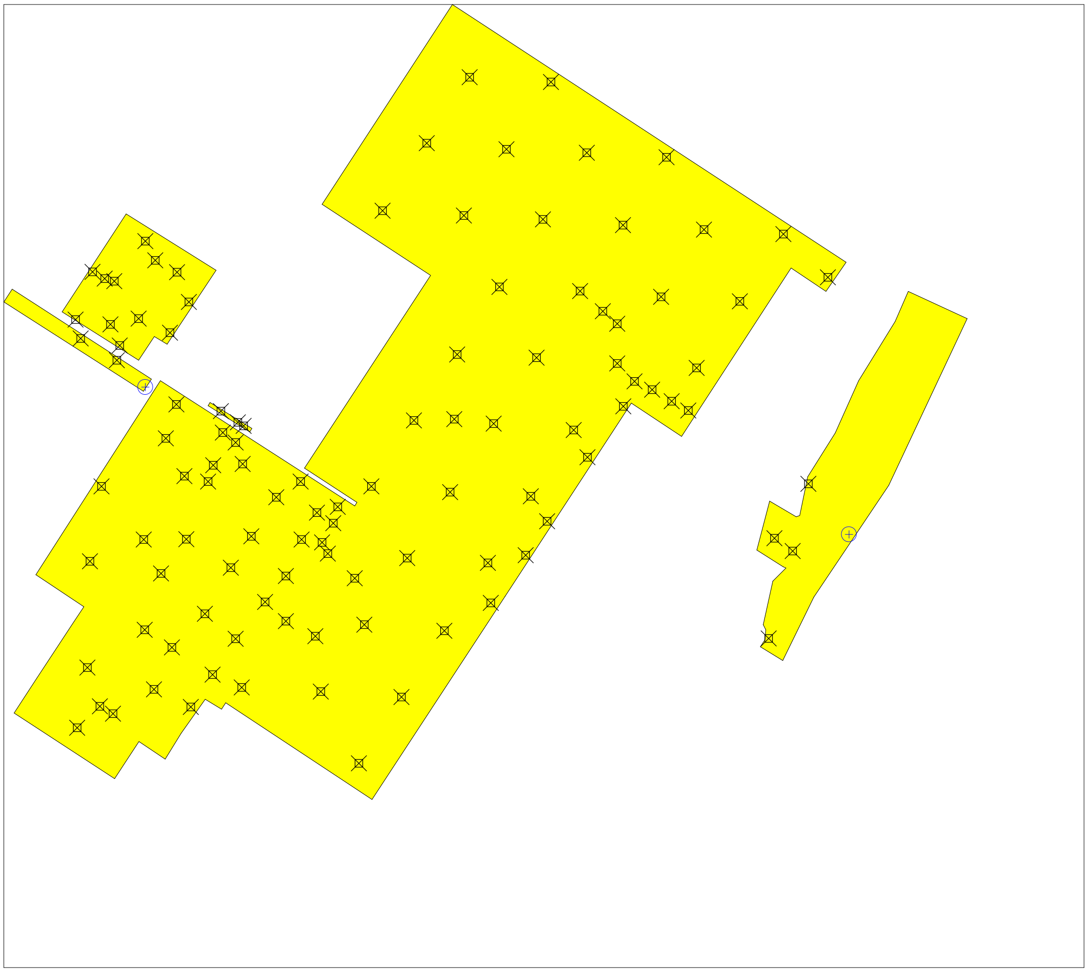
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

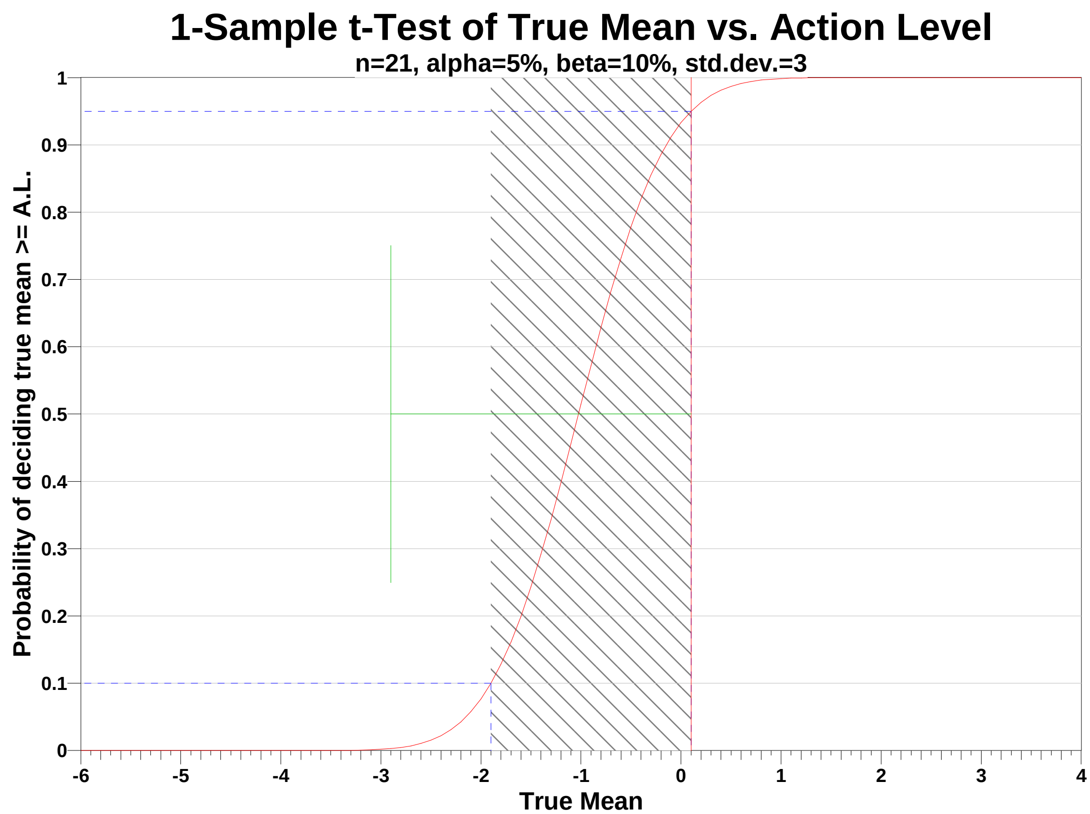
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=0.1		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	3895984	973997	3082986	770748	2588149	647038
	β=10	3082987	770748	2365020	591256	1934301	483576
	β=15	2588150	647039	1934301	483576	1546841	386711
LBGR=80	β=5	973997	243501	770748	192688	647038	161760
	β=10	770748	192688	591256	147815	483576	120895
	β=15	647039	161761	483576	120895	386711	96679
LBGR=70	β=5	432889	108224	342555	85640	287573	71894

	β=10	342556	85640	262781	65696	214923	53732
	β=15	287574	71895	214923	53732	171872	42969

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n				110				
Min				0				
Max				3.97				
Range				3.97				
Mean				0.18668				
Median				0.07925				
Variance				0.1938				
StdDev				0.44023				
Std Error				0.041974				
Skewness				6.7771				
Interquartile Range				0.0245				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.00682	0.06855	0.0711	0.074	0.07925	0.0985	0.401	0.729	3.761

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	8.569	3.405	Yes

The test statistic 8.569 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	3.97

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.4027
Lilliefors 5% Critical Value	0.08526

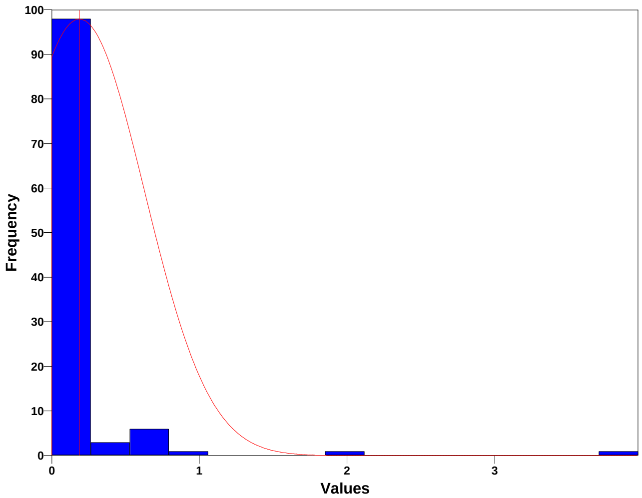
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

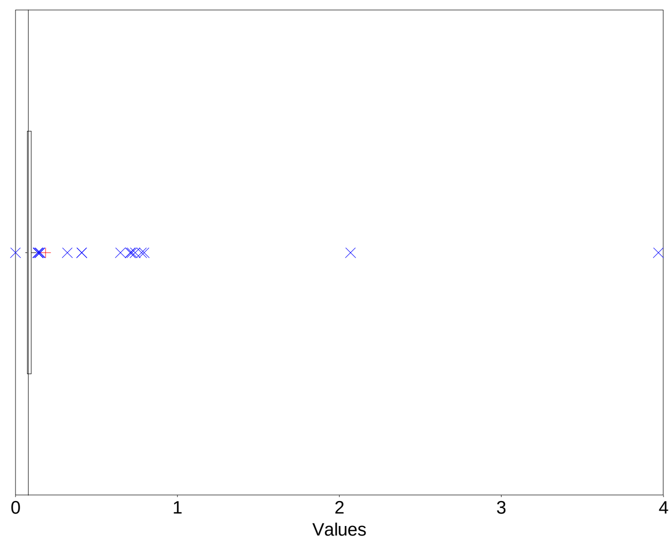
Data Plots
Graphical displays of the data are shown below.

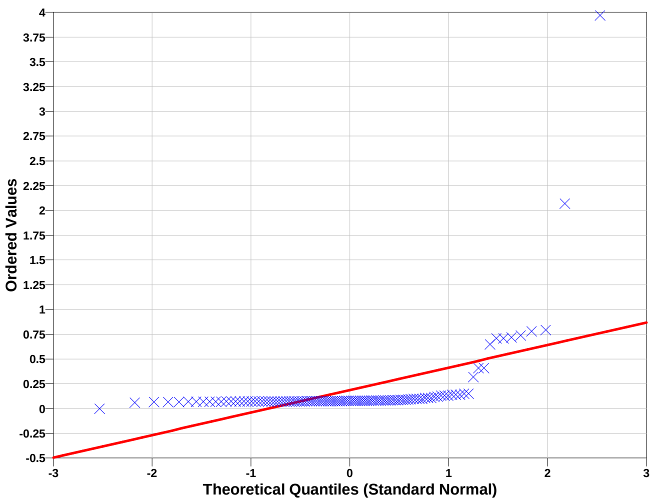
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25th and 75th percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The pth quantile of a distribution of data is the data value, x_n, for which a fraction p of the distribution is less than x_n. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.4219
Lilliefors 5% Critical Value	0.08448

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.2563

95% Non-Parametric (Chebyshev) UCL	0.3696

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.3696) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=110 data,
AL is the action level or threshold (0.1),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=109 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
2.0651	1.659	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
83	63	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

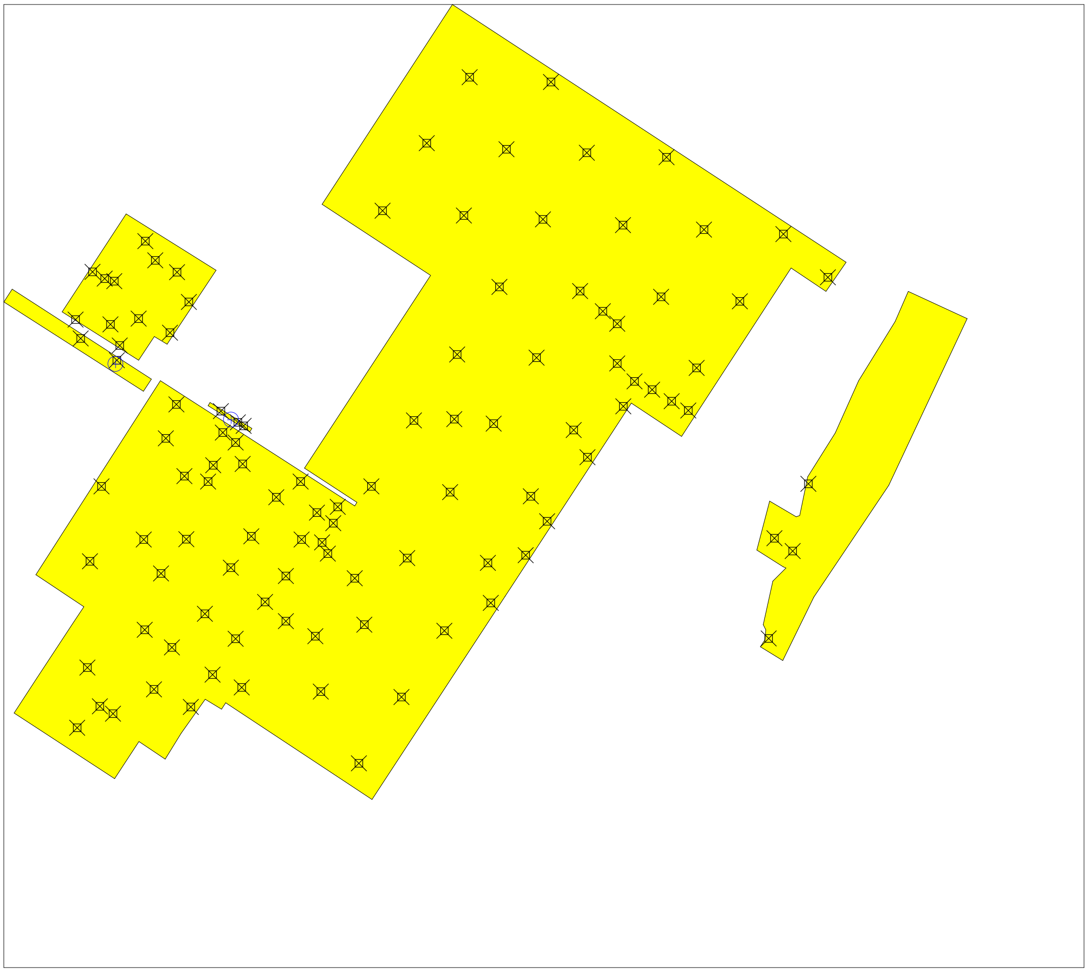
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

The values of these inputs that result in the calculated number of sampling locations are:

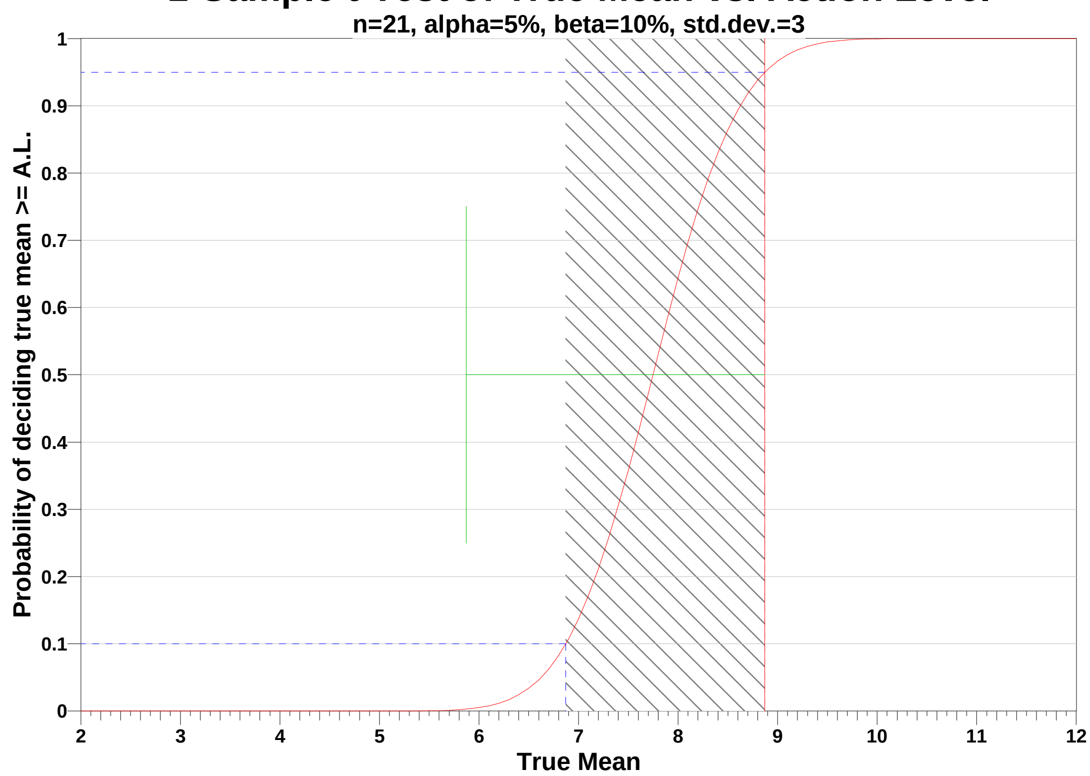
Analyte	n	Parameter					
		S	**Δ**	**α**	**β**	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.

1-Sample t-Test of True Mean vs. Action Level



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=8.87154		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	497	126	393	99	330	83
	β=10	394	100	302	76	247	62
	β=15	331	84	247	63	198	50
LBGR=80	**β=5**	126	33	99	26	83	22
	β=10	100	26	76	20	62	16
	β=15	84	22	63	17	50	13
LBGR=70	**β=5**	57	16	45	12	38	10

	β=10	45	13	35	10	28	8
	β=15	38	11	29	8	23	6

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	112							
Min	0							
Max	3.97							
Range	3.97							
Mean	0.18468							
Median	0.07925							
Variance	0.19054							
StdDev	0.43651							
Std Error	0.041246							
Skewness	6.8383							
Interquartile Range	0.024							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.00806	0.068	0.0703	0.074	0.07925	0.098	0.383	0.727	3.723

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	8.633	3.411	Yes

The test statistic 8.633 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	3.97

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.3982
Lilliefors 5% Critical Value	0.08448

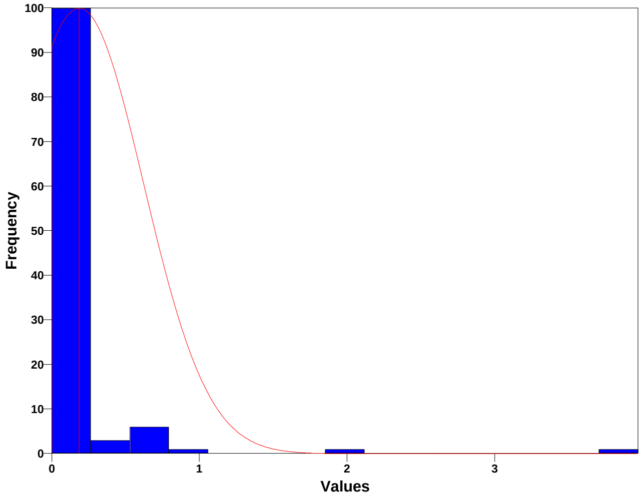
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

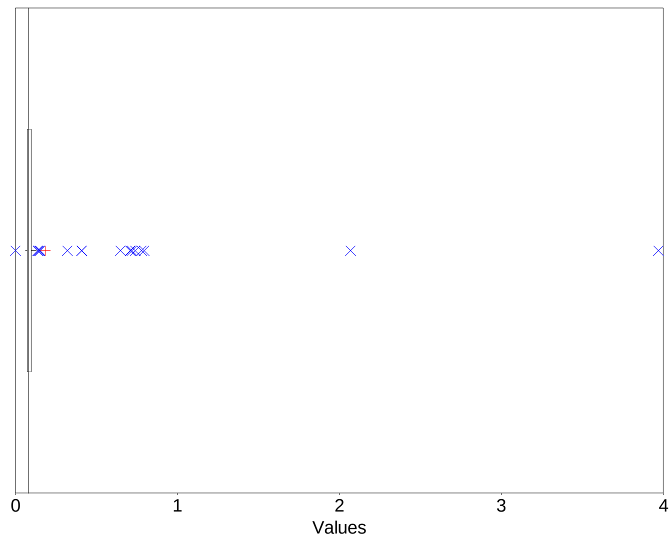
Data Plots
Graphical displays of the data are shown below.

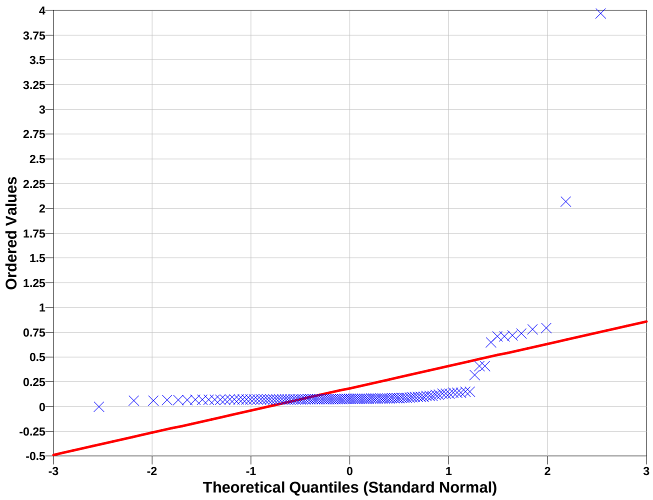
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests
A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.4222
Lilliefors 5% Critical Value	0.08372

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean
Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.2531

95% Non-Parametric (Chebyshev) UCL	0.3645

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.3645) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=112 data,
AL is the action level or threshold (8.87154),
SE is the standard error = (standard deviation) / (square root of n).

This *t* was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=111 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-210.61	1.6587	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
112	65	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

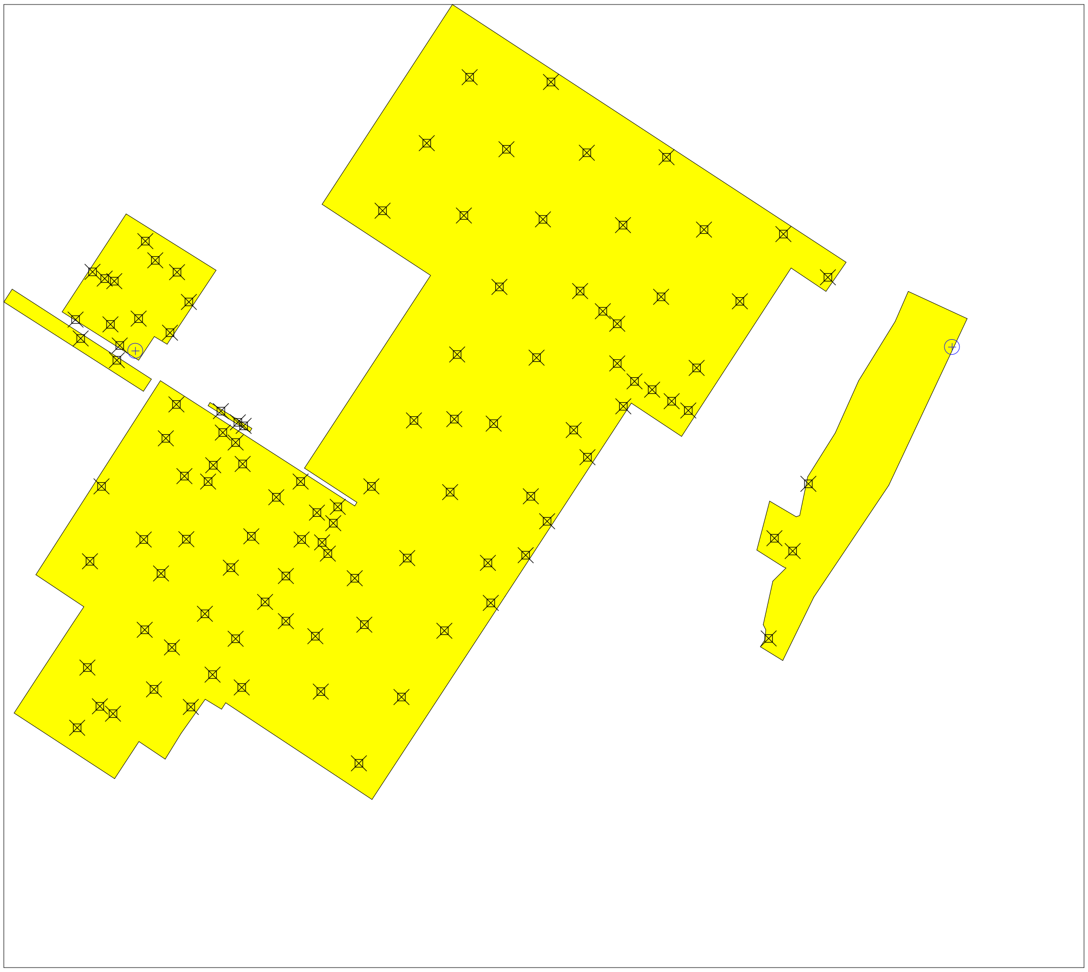
Random sampling locations for comparing a mean with a fixed threshold (parametric)

Summary
This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where
n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

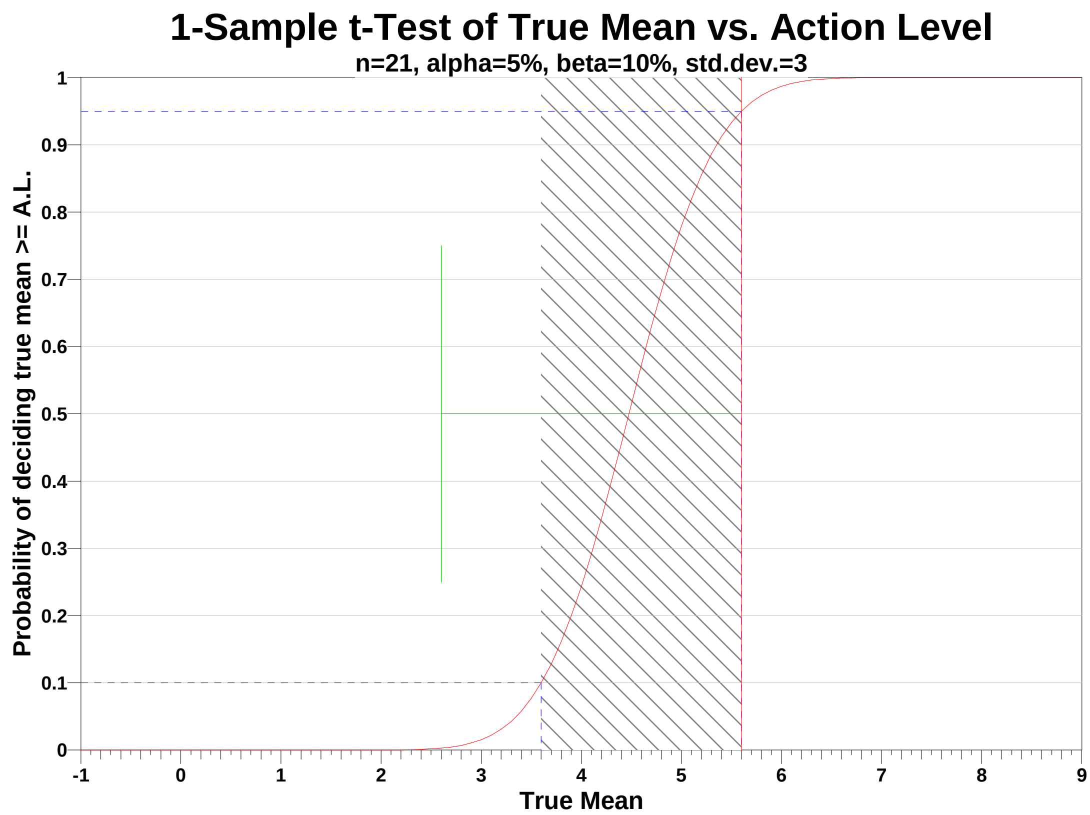
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	n	Parameter					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=5.6		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	1244	312	984	247	826	207
	β=10	985	248	755	190	618	155
	β=15	827	208	618	156	494	124
LBGR=80	**β=5**	312	79	247	63	207	53
	β=10	248	63	190	48	155	40
	β=15	208	53	156	40	124	32
LBGR=70	**β=5**	140	36	111	29	93	24

	β=10	111	29	85	22	70	18
	β=15	94	25	70	18	56	15

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n				110				
Min				0				
Max				3.97				
Range				3.97				
Mean				0.18668				
Median				0.07925				
Variance				0.1938				
StdDev				0.44023				
Std Error				0.041974				
Skewness				6.7771				
Interquartile Range				0.0245				
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.00682	0.06855	0.0711	0.074	0.07925	0.0985	0.401	0.729	3.761

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	8.594	3.408	Yes

The test statistic 8.594 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	3.97

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.3982
Lilliefors 5% Critical Value	0.08486

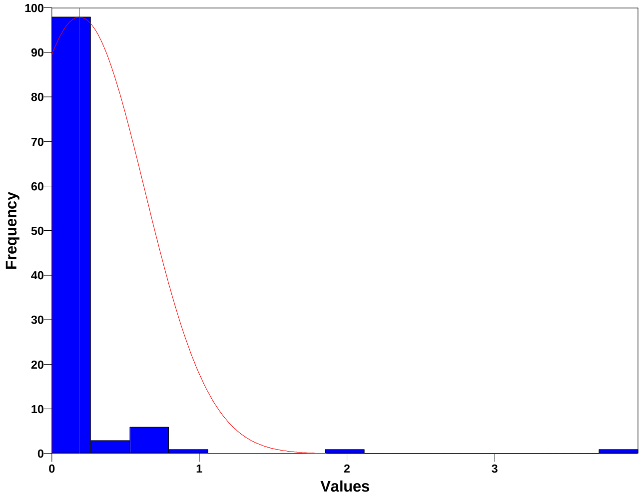
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

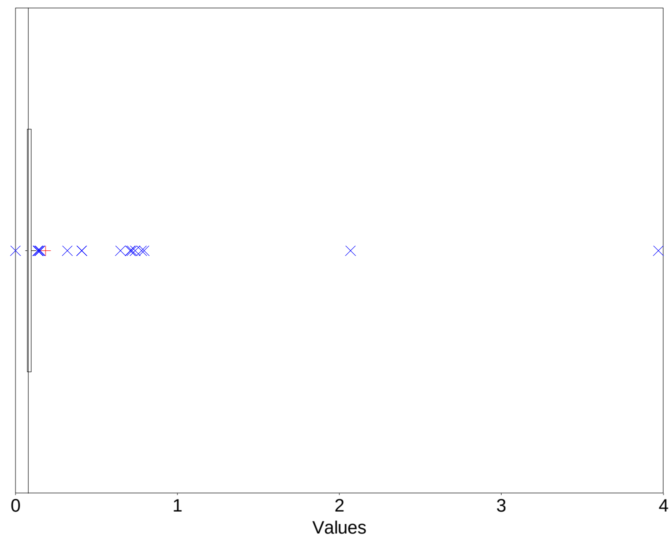
Data Plots
Graphical displays of the data are shown below.

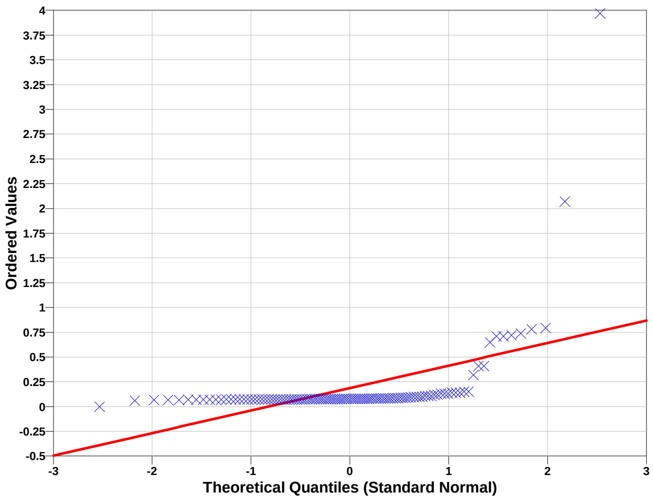
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.4219
Lilliefors 5% Critical Value	0.08448

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	0.2563

95% Non-Parametric (Chebyshev) UCL	0.3696

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (0.3696) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=110 data,
AL is the action level or threshold (5.6),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=109 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-128.97	1.659	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
110	64	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

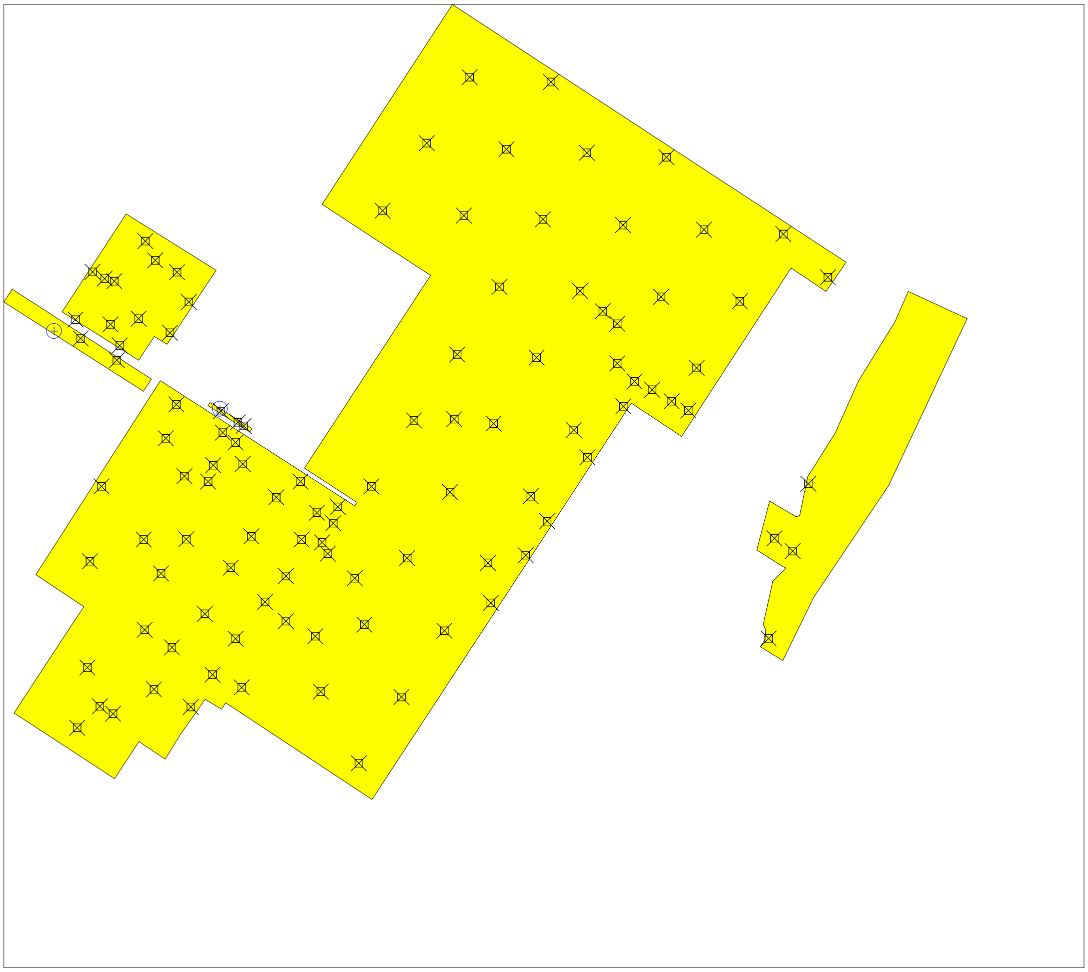
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n	is the number of samples,
S	is the estimated standard deviation of the measured values including analytical error,
Δ	is the width of the gray region,
α	is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β	is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$	is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$	is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

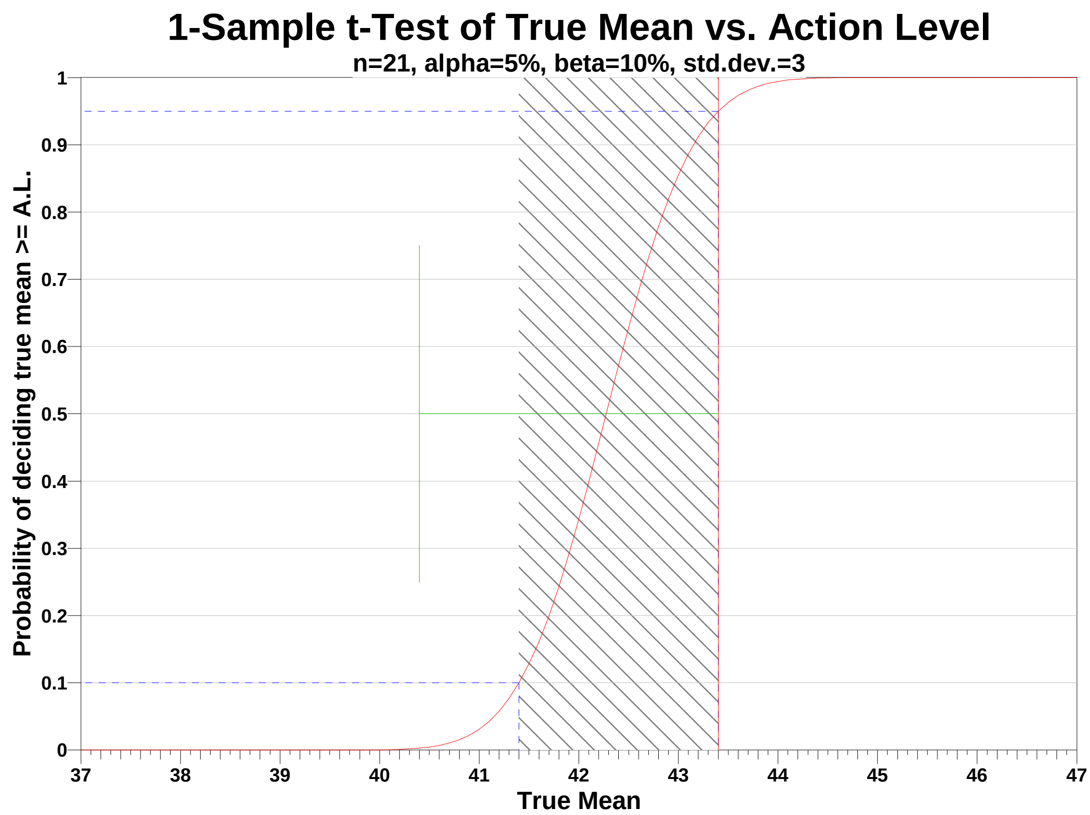
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	**n**	**Parameter**					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=43.4		α=5		α=10		α=15	
		s=6	s=3	s=6	s=3	s=6	s=3
LBGR=90	β=5	23	7	18	5	15	4
	β=10	18	6	14	4	11	4
	β=15	16	5	12	4	9	3
LBGR=80	β=5	7	3	5	2	4	2
	β=10	6	3	4	2	4	2
	β=15	5	3	4	2	3	2
LBGR=70	β=5	4	2	3	2	3	1

	β=10	4	2	3	2	2	1
	β=15	3	2	2	2	2	1

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	110							
Min	0							
Max	270							
Range	270							
Mean	7.8219							
Median	3.8							
Variance	664.6							
StdDev	25.78							
Std Error	2.458							
Skewness	9.8371							
Interquartile Range	4.95							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0.05555	1.019	1.305	2.075	3.8	7.025	13.2	16.25	243.6

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	10.12	3.405	Yes

The test statistic 10.12 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	270

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.2254
Lilliefors 5% Critical Value	0.08526

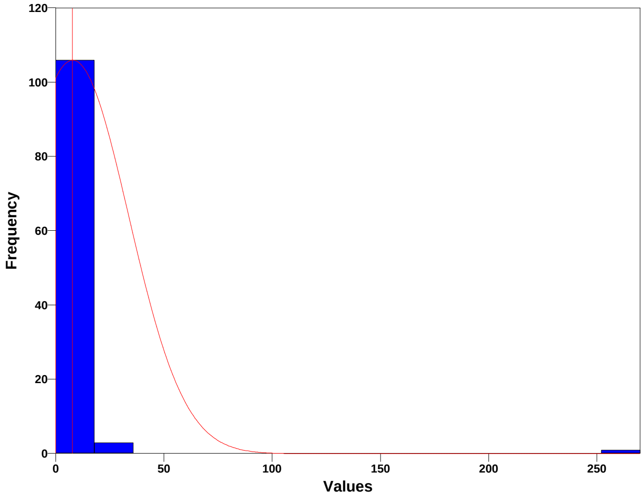
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

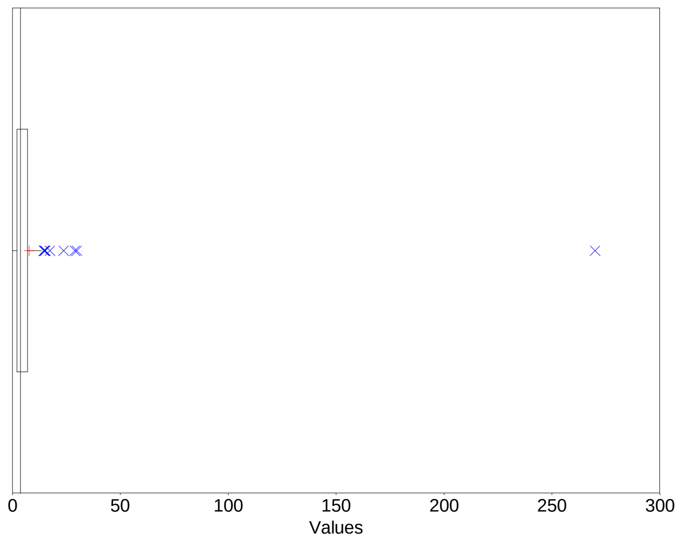
Data Plots
Graphical displays of the data are shown below.

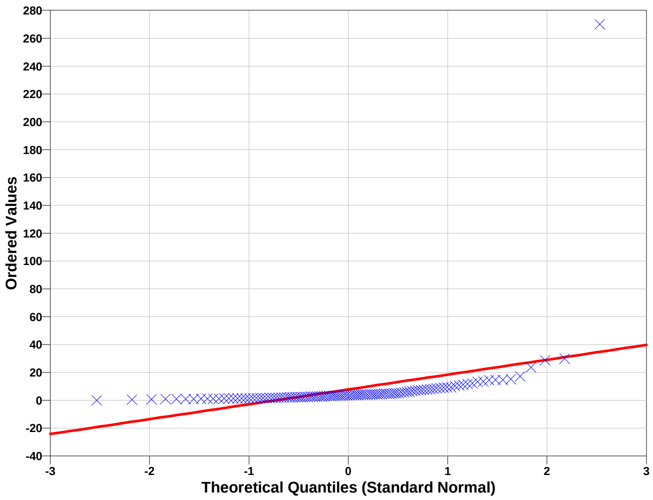
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.3808
Lilliefors 5% Critical Value	0.08448

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	11.9

95% Non-Parametric (Chebyshev) UCL	18.54

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (18.54) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value *t* was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=110 data,
AL is the action level or threshold (43.4),
SE is the standard error = (standard deviation) / (square root of n).

This *t* was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=109 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-14.474	1.659	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
109	64	Reject

This report was automatically produced* by Visual Sample Plan (VSP) software version 5.000.
Software and documentation available at http://dqo.pnl.gov/vsp

* - The report contents may have been modified or reformatted by end-user of software.

Random sampling locations for comparing a mean with a fixed threshold (parametric)

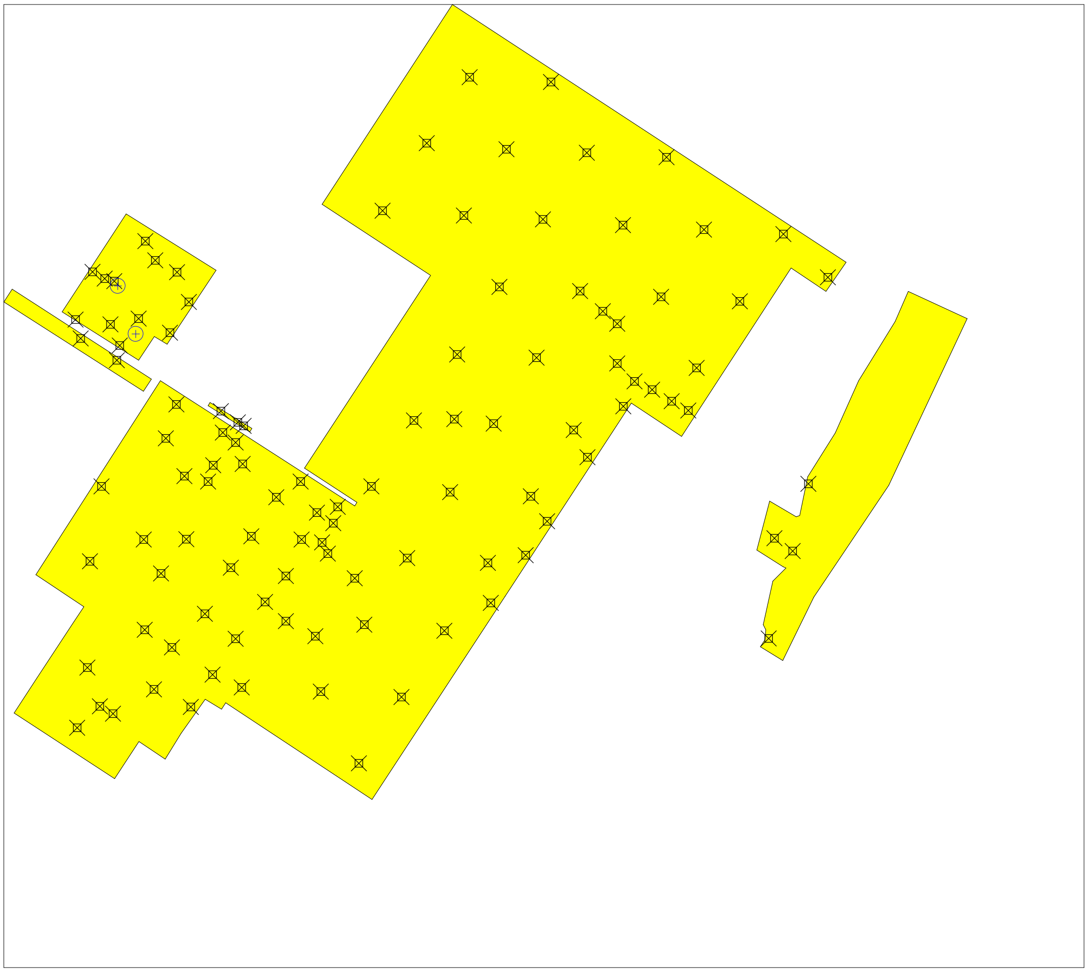
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean to a fixed threshold
Type of Sampling Design	Parametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The mean value at the site exceeds the threshold
Formula for calculating number of sampling locations	Student's t-test
Calculated total number of samples	21
Number of samples on map [a]	109
Number of selected sample areas [b]	5
Specified sampling area [c]	941426.62 m^2
Total cost of sampling [d]	$11,500.00

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Including measurement analyses and fixed overhead costs. See the Cost of Sampling section for an explanation of the costs presented here.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the mean value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the mean value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative hypothesis, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A parametric random sampling approach was used to determine the number of samples and to specify sampling locations. A parametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that parametric assumptions are reasonable. These assumptions will be examined in post-sampling data analysis.

Both parametric and non-parametric approaches rely on assumptions about the population. However, non-parametric approaches typically require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than the number of samples required by non-parametric approaches.

Locating the sample points randomly provides data that are separated by many distances, whereas systematic samples are all equidistant apart. Therefore, random sampling provides more information about the spatial structure of the potential contamination than systematic sampling does. As with systematic sampling, random sampling also provides information regarding the mean value, but there is the possibility that areas of the site will not be represented with the same frequency as if uniform grid sampling were performed.

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Student's t-test. For this site, the null hypothesis is rejected in favor of the alternative hypothesis if the sample mean is sufficiently smaller than the threshold. The number of

samples to collect is calculated so that 1) there will be a high probability ($1-\beta$) of rejecting the null hypothesis if the alternative hypothesis is true and 2) a low probability (α) of rejecting the null hypothesis if the null hypothesis is true.

The formula used to calculate the number of samples is:

$$n = \frac{S^2}{\Delta^2}\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 + 0.5Z_{1-\alpha}^2$$

where

n is the number of samples,
S is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site mean is less than the threshold,
β is the acceptable probability of incorrectly concluding the site mean exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

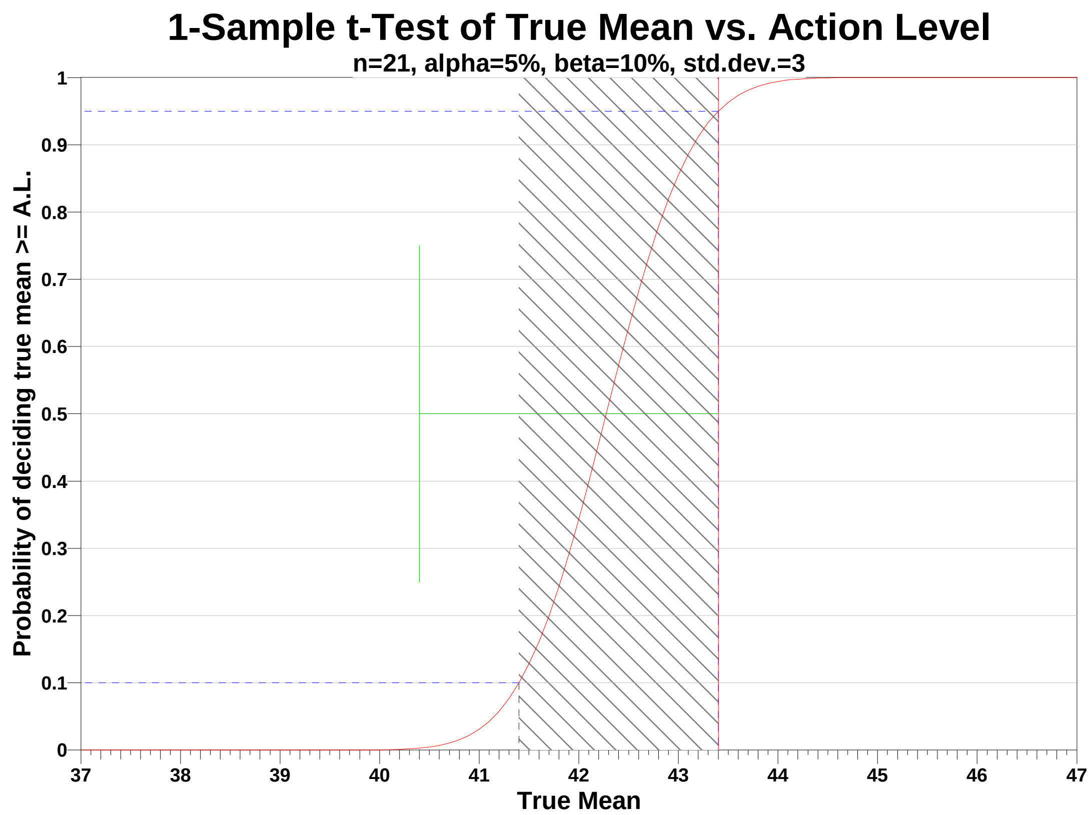
The values of these inputs that result in the calculated number of sampling locations are:

Analyte	**n**	**Parameter**					
		S	Δ	α	β	$Z_{1-\alpha}$ [a]	$Z_{1-\beta}$ [b]
	21	3	2	0.05	0.1	1.64485	1.28155

[a] This value is automatically calculated by VSP based upon the user defined value of α.
[b] This value is automatically calculated by VSP based upon the user defined value of β.

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true mean values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the sample mean is normally distributed (this happens if the data are roughly symmetric and the sample size is 30 or more; for skewed data sets, additional samples are required for the sample mean to be normally distributed),
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis
The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level and examining the resulting changes in the number of samples. The following table shows the results of this analysis.

Number of Samples							
AL=43.4		**α=5**		**α=10**		**α=15**	
		s=6	**s=3**	**s=6**	**s=3**	**s=6**	**s=3**
LBGR=90	**β=5**	23	7	18	5	15	4
	β=10	18	6	14	4	11	4
	β=15	16	5	12	4	9	3
LBGR=80	**β=5**	7	3	5	2	4	2
	β=10	6	3	4	2	4	2
	β=15	5	3	4	2	3	2
LBGR=70	**β=5**	4	2	3	2	3	1

	β=10	4	2	3	2	2	1
	β=15	3	2	2	2	2	1

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Cost of Sampling
The total cost of the completed sampling program depends on several cost inputs, some of which are fixed, and others that are based on the number of samples collected and measured. Based on the numbers of samples determined above, the estimated total cost of sampling and analysis at this site is $11,500.00, which averages out to a per sample cost of $547.62. The following table summarizes the inputs and resulting cost estimates.

COST INFORMATION			
Cost Details	**Per Analysis**	**Per Sample**	**21 Samples**
Field collection costs		$100.00	$2,100.00
Analytical costs	$400.00	$400.00	$8,400.00
Sum of Field & Analytical costs		**$500.00**	**$10,500.00**
Fixed planning and validation costs			$1,000.00
Total cost			**$11,500.00**

Data Analysis

SUMMARY STATISTICS								
n	112							
Min	0							
Max	270							
Range	270							
Mean	7.7916							
Median	3.8							
Variance	653.35							
StdDev	25.561							
Std Error	2.4153							
Skewness	9.9105							
Interquartile Range	5.25							
Percentiles								
1%	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**
0	0.913	1.265	2.025	3.8	7.275	12.99	16.03	238.8

Outlier Test
Rosner's test for multiple outliers was performed to test whether the most extreme value is a statistical outlier. The test was conducted at the 5% significance level.

Data should not be excluded from analysis solely on the basis of the results of this or any other statistical test. If any

values are flagged as possible outliers, further investigation is recommended to determine whether there is a plausible explanation that justifies removing or replacing them.

In using Rosner's test to detect up to 1 outlier, a test statistic R_1 is calculated, and compared with a critical value C_1 to test the hypothesis that there is one outlier in the data.

ROSNER'S OUTLIER TEST			
k	**Test Statistic R_k**	**5% Critical Value C_k**	**Significant?**
1	10.21	3.411	Yes

The test statistic 10.21 exceeded the corresponding critical value, therefore that test is significant and we conclude that the most extreme value is an outlier at the 5% significance level.

SUSPECTED OUTLIERS	
1	270

A normal distribution test indicated that the data do not appear to be normally distributed, so further investigation is recommended before using the results of this test. Because Rosner's test can be used only when the data without the suspected outlier are approximately normally distributed, a Lilliefors test for normality was performed at a 5% significance level.

NORMAL DISTRIBUTION TEST (excluding outliers)	
Lilliefors Test Statistic	0.2245
Lilliefors 5% Critical Value	0.08448

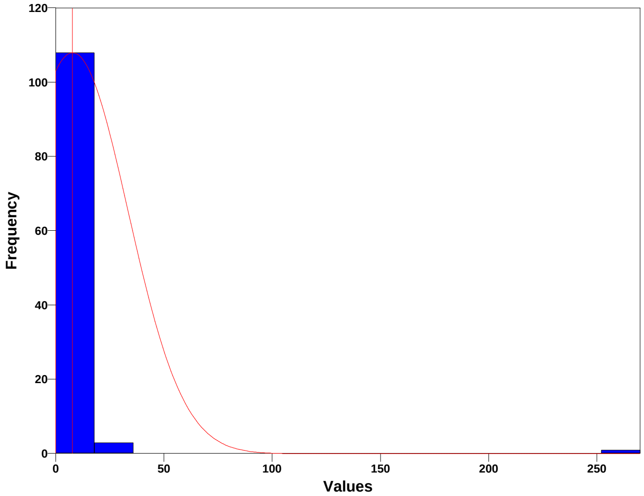
The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so the test rejects the hypothesis that the data are normal and concludes that the data, excluding the most extreme value, do not appear to follow a normal distribution at the 5% level of significance. Rosner's test may not be appropriate if the assumption of normally distributed data is not justified for this data set. Examine the Q-Q plot displayed below to further assess the normality of the data.

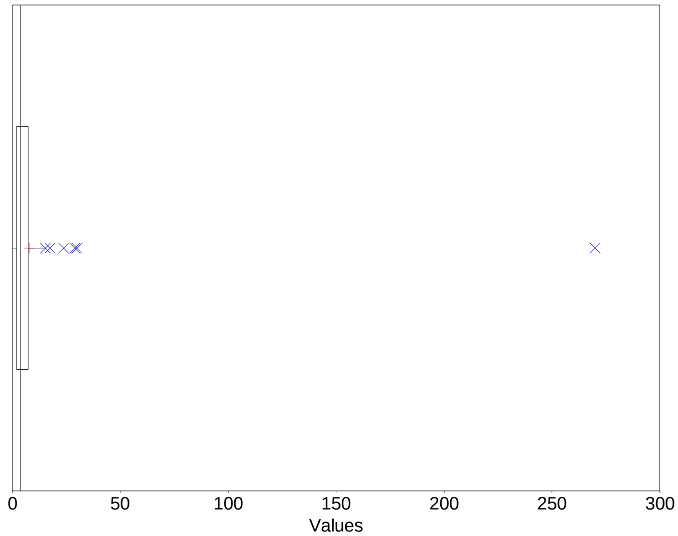
Data Plots
Graphical displays of the data are shown below.

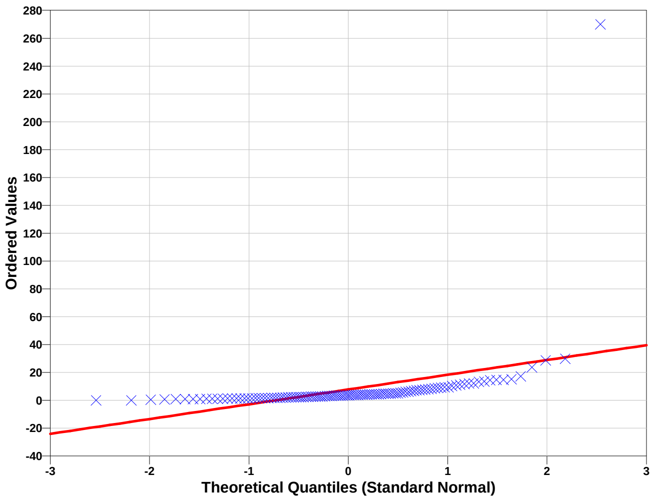
The Histogram is a plot of the fraction of the n observed data that fall within specified data “bins.” A histogram is generated by dividing the x axis (range of the observed data values) into "bins" and displaying the number of data in each bin as the height of a bar for the bin. The area of the bar is the fraction of the n data values that lie within the bin. The sum of the fractions for all bins equals one. A histogram is used to assess how the n data are distributed (spread) over their range of values. If the histogram is more or less symmetric and bell shaped, then the data may be normally distributed.

The Box and Whiskers plot is composed of a central box divided by a line, and with two lines extending out from the box, called the "whiskers". The line through the box is drawn at the median of the n data observed. The two ends of the box represent the 25^{th} and 75^{th} percentiles of the n data values, which are also called the lower and upper quartiles, respectively, of the data set. The sample mean (mean of the n data) is shown as a "+" sign. The upper whisker extends to the largest data value that is less than the upper quartile plus 1.5 times the interquartile range (upper quartile minus the lower quartile). The lower whisker extends to the smallest data value that is greater than the lower quartile minus 1.5 times the interquartile range. Extreme data values (greater or smaller than the ends of the whiskers) are plotted individually as blue Xs. A Box and Whiskers plot is used to assess the symmetry of the distribution of the data set. If the distribution is symmetrical, the box is divided into two equal halves by the median, the whiskers will be the same length, and the number of extreme data points will be distributed equally on either end of the plot.

The Q-Q plot graphs the quantiles of a set of n data against the quantiles of a specific distribution. We show here only the Q-Q plot for an assumed normal distribution. The p^{th} quantile of a distribution of data is the data value, x_p, for which a fraction p of the distribution is less than x_p. If the data plotted on the normal distribution Q-Q plot closely follow a straight line, even at the ends of the line, then the data may be assumed to be normally distributed. If the data points deviate substantially from a linear line, then the data are not normally distributed.







For more information on these plots consult Guidance for Data Quality Assessment, EPA QA/G-9, pgs 2.3-1 through 2.3-12. (http://www.epa.gov/quality/qa-docs.html).

Tests

A goodness-of-fit test was performed to test whether the data set had been drawn from an underlying normal distribution. The Lilliefors test was used to test the null hypothesis that the data are normally distributed. The test was conducted at the 5% significance level, i.e., the probability the test incorrectly rejects the null hypothesis was set at 0.05.

NORMAL DISTRIBUTION TEST	
Lilliefors Test Statistic	0.3802
Lilliefors 5% Critical Value	0.08372

The calculated Lilliefors test statistic exceeds the 5% Lilliefors critical value, so we can reject the hypothesis that the data are normal, or in other words the data do not appear to follow a normal distribution at the 5% level of significance. The Q-Q plot displayed above should be used to further assess the normality of the data.

Upper Confidence Limit on the True Mean

Two methods were used to compute the upper confidence limit (UCL) on the mean. The first is a parametric method that assumes a normal distribution. The second is the Chebyshev method, which requires no distributional assumption.

UCLs ON THE MEAN	
95% Parametric UCL	11.8

95% Non-Parametric (Chebyshev) UCL	18.32

Because the data do not appear to be normally distributed according to the goodness-of-fit test performed above, the non-parametric UCL (18.32) may be a more accurate upper confidence limit on the true mean.

One-Sample t-Test
A one-sample t-test was performed to compare the sample mean to the action level. The null hypothesis used is that the true mean equals or exceeds the action level (AL). The t-test was conducted at the 5% significance level. The sample value t was computed using the following equation:

$$t = \frac{\bar{x} - AL}{SE}$$

where
x is the sample mean of the n=112 data,
AL is the action level or threshold (43.4),
SE is the standard error = (standard deviation) / (square root of n).

This t was then compared with the critical value $t_{0.95}$, where $t_{0.95}$ is the value of the t distribution with n-1=111 degrees of freedom for which the proportion of the distribution to the left of $t_{0.95}$ is 0.95. The null hypothesis will be rejected if $t < -t_{0.95}$.

ONE-SAMPLE t-TEST		
t-statistic	**Critical Value $t_{0.95}$**	**Null Hypothesis**
-14.743	1.6587	Reject

The test rejected the null hypothesis that the mean value at the site exceeds the threshold, therefore conclude the true mean is less than the threshold.

Because the data do not appear to be normally distributed, the MARSSIM Sign Test might be preferred over the One Sample t-Test. The following table represents the results of the MARSSIM Sign Test using the current data:

MARSSIM Sign Test		
Test Statistic (S+)	**95% Critical Value**	**Null Hypothesis**
111	65	Reject